

Db 10 NHINCYLYFII 21

RESULT 140

US-10-424-599-232693

; Sequence 232693, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 232693

; LENGTH: 48

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_52146C.1.pep

US-10-424-599-232693

Query Match 29.9%; Score 29; DB 15; Length 48;

Best Local Similarity 44.4%; Pred. No. 2.2e+03;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLNSKI 9

Db 22 QPQHLSSSV 30

RESULT 141

US-10-424-599-155980

; Sequence 155980, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 155980

; LENGTH: 49

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_111870C.1.pep

US-10-424-599-155980

Query Match 29.9%; Score 29; DB 15; Length 49;

Best Local Similarity 62.5%; Pred. No. 2.2e+03;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 HNSKIAP 11

Db 21 HLNHKVFF 28

RESULT 142

US-10-424-599-270499

; Sequence 270499, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

US-10-424-599-270499

Query Match 29.9%; Score 29; DB 10; Length 50;

Best Local Similarity 38.5%; Pred. No. 2.3e+03;

Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PNHLNSKIAPKIV 14

Db 30 PPHVNMSELVKII 42

US-10-424-599-270499

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 270499

; LENGTH: 49

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_86279C.1.pep

US-10-424-599-270499

Query Match 29.9%; Score 29; DB 15; Length 49;

Best Local Similarity 31.2%; Pred. No. 2.2e+03;

Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHLNSKIAPKIVSOEP 18

Db 29 NNLLNLFQSFINMEP 44

RESULT 143

US-09-820-649-221

; Sequence 221, Application US/09820649

; Publication No. US20030199683A1

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: 83 Human Secreted Proteins

; FILE REFERENCE: P2012P1

; CURRENT APPLICATION NUMBER: US/09/820,649

; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: US/09/236,557

; PRIOR FILING DATE: 1999-01-26

; PRIOR APPLICATION NUMBER: PCT/US98/15949

; PRIOR FILING DATE: 1998-07-29

; PRIOR APPLICATION NUMBER: 60/054,212

; PRIOR FILING DATE: 1997-07-30

; PRIOR APPLICATION NUMBER: 60/054,209

; PRIOR FILING DATE: 1997-07-30

; PRIOR APPLICATION NUMBER: 60/054,234

; PRIOR FILING DATE: 1997-07-30

; PRIOR APPLICATION NUMBER: 60/054,218

; PRIOR FILING DATE: 1997-07-30

; PRIOR APPLICATION NUMBER: 60/054,214

; PRIOR FILING DATE: 1997-07-30

; PRIOR APPLICATION NUMBER: 60/054,236

; PRIOR FILING DATE: 1997-07-30

; PRIOR APPLICATION NUMBER: 60/054,215

; PRIOR FILING DATE: 1997-07-30

; PRIOR APPLICATION NUMBER: 60/054,211

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 353

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 221

; LENGTH: 50

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-820-649-221

Query Match 29.9%; Score 29; DB 10; Length 50;

Best Local Similarity 38.5%; Pred. No. 2.3e+03;

Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PNHLNSKIAPKIV 14

Db 30 PPHVNMSELVKII 42

Db 10 NHINCYLYFII 21

RESULT 140

US-10-424-599-232693

; Sequence 232693, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 232693

; LENGTH: 48

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_52146C.1.pep

US-10-424-599-232693

Query Match 29.9%; Score 29; DB 15; Length 48;

Best Local Similarity 44.4%; Pred. No. 2.2e+03;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLNSKI 9

Db 22 QPQHLSSSV 30

RESULT 141

US-10-424-599-155980

; Sequence 155980, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 155980

; LENGTH: 49

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_111870C.1.pep

US-10-424-599-155980

Query Match 29.9%; Score 29; DB 15; Length 49;

Best Local Similarity 62.5%; Pred. No. 2.2e+03;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 HNSKIAP 11

Db 21 HLNHKVFF 28

RESULT 142

US-10-424-599-270499

; Sequence 270499, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

RESULT 144

```

US-10-160-162-221
; Sequence 221, Application US/10160162
; Publication No. US20030166541A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 83 Human Secreted Proteins
; FILE REFERENCE: PZ012P2
; CURRENT APPLICATION NUMBER: US/10/160,162
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/295,558
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 09/236,557
; PRIOR FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: PCT/US98/15949
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,212
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,209
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,234
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,218
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,214
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,236
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,215
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,211
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,217
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,213
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/055,968
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,969
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,972
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,561
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,534
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,729
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,543
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,727
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,554
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,730
; PRIOR FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 221
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-160-162-221

```

RESULT 145

```

US-09-864-761-47206
/ Sequence 47206, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aemica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/532,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 47206
/ LENGTH: 30
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC005690.7
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.78
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.6
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.54
US-09-864-761-47206

```

```
RESULT 146
US-10-424-599-147555
; Sequence 147555, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 147555
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_104260C.1.pep
US-10-424-599-147555
```

```
Query Match          29.4%; Score 28.5; DB 15; Length 33;
Best Local Similarity 23.5%; Pred. No. 1.7e+03;
Matches 4; Conservative 8; Mismatches 4; Indels 1; Gaps 1;
```

```
QY      2 PNHLNSKIAFKIVSQEP 18
|||: : : : : : :
Db      11 PNHGASLEY-VLNRPV 26

RESULT 147
US-10-424-599-210402
; Sequence 210402, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 210402
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_3201C.1.pep
US-10-424-599-210402
```

```
Query Match          29.4%; Score 28.5; DB 15; Length 39;
Best Local Similarity 63.6%; Pred. No. 2.1e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
```

```
QY      1 EPNHLNSKIAF 11
|| : : : : :
Db      20 EP-YLNKKITF 29

RESULT 148
US-10-437-963-154010
; Sequence 154010, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154010
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53910C.1.pep
US-10-437-963-154010
```

```
Query Match          29.4%; Score 28.5; DB 16; Length 43;
Best Local Similarity 35.0%; Pred. No. 2.3e+03;
Matches 7; Conservative 4; Mismatches 8; Indels 1; Gaps 1;
```

```
QY      1 EPNHLNSKIA-FKIVSQEPA 19
|||: : : : : :
Db      23 EINNKGKCEDFVLGQRP 42
```

```
RESULT 149
US-09-864-761-44631
; Sequence 44631, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeowica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
```


; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 44631
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC018528.3
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.75
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.84
 ; US-09-864-761-44631

Query Match 28.9%; Score 28; DB 9; Length 19;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 NSKIAFKIVSQE 17

Db 6 NYKLSVKFSQEQ 17

RESULT 150

; US-09-982-172-55
 ; Sequence 55, Application US/09982172
 ; Patent No. US20020137119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Emil Israel Katz
 ; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
 ; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT
 ; TITLE OF INVENTION: UTILIZING EACH
 ; FILE REFERENCE: 01/22283
 ; CURRENT APPLICATION NUMBER: US/09/982,172
 ; CURRENT FILING DATE: 2001-10-19
 ; NUMBER OF SEQ ID NOS: 253
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 55
 ; LENGTH: 22
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Computer generated synthetic peptide
 ; US-09-982-172-55

Query Match 28.9%; Score 28; DB 9; Length 22;
 Best Local Similarity 83.3%; Pred. No. 1.3e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 13 IVSQEQP 18

Db 5 VVSQEQP 10

RESULT 151

; US-10-242-056-70
 ; Sequence 70, Application US/10242056
 ; Publication No. US20030113323A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ensign, Gerald C
 ; APPLICANT: Bowen, David J
 ; APPLICANT: Petell, James
 ; APPLICANT: Fatig, Raymond
 ; APPLICANT: Schoonover, Sue
 ; APPLICANT: ffrench-Constant, Richard
 ; APPLICANT: Orr, Gregory L

; APPLICANT: Merlo, Donald J
 ; APPLICANT: Roberts, Jean L
 ; APPLICANT: Rocheleau, Thomas A
 ; TITLE OF INVENTION: Insecticidal Protein Toxins from
 ; TITLE OF INVENTION: Photorhabdus
 ; NUMBER OF SEQUENCES: 88
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DowElanco
 ; STREET: 9330 Zionsville Road
 ; CITY: Indianapolis
 ; STATE: IN
 ; COUNTRY: US
 ; ZIP: 46268
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/242,056
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/063,615
 ; FILING DATE: 18-MAY-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/395,497
 ; FILING DATE: 28-FEB-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/007,255
 ; FILING DATE: 06-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/608,423
 ; FILING DATE: 28-FEB-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/705,484
 ; FILING DATE: 28-AUG-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/743,699
 ; FILING DATE: 06-NOV-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Borucki, Andrea T.
 ; REGISTRATION NUMBER: 33651
 ; REFERENCE/DOCKET NUMBER: 50301E
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 317-337-4846
 ; TELEFAX: 317-337-4847
 ; INFORMATION FOR SEQ ID NO: 70: Tcdali-syn
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 22 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal
 ; US-10-242-056-70

Query Match 28.9%; Score 28; DB 14; Length 22;
 Best Local Similarity 40.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 PNHLSKIAFKIVSQ 16

Db 3 PNNSSNKLMPYVYQ 17

RESULT 152

; US-10-262-794A-70
 ; Sequence 70, Application US/10262794A
 ; Publication No. US20030207806A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ensign, Gerald C
 ; APPLICANT: Bowen, David J

Db 3 PNNSSNKLMPFVYQ 17

RESULT 153

US-09-864-761-34851

; Sequence 34851, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Petr. Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Asomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 34851

; LENGTH: 23

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO A1109933.5

; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.7

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 15

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.4

; OTHER INFORMATION: EST_HUMAN HIT: AV694621.1, EVALUE 8.20e+00

US-09-864-761-34851

APPLICANT: Petell, James

APPLICANT: Fatig, Raymond

APPLICANT: Schoonover, Sue

APPLICANT: French-Constant, Richard

APPLICANT: Rocheleau, Thomas A.

APPLICANT: Blackburn, Michael B.

APPLICANT: Hey, Timothy D.

APPLICANT: Merlo, Donald J.

APPLICANT: Orr, Gregory L.

APPLICANT: Roberts, Jean L.

APPLICANT: Strickland, James A.

APPLICANT: Guo, Lining

APPLICANT: Ciche, Todd A.

APPLICANT: Sukhapinda, Kitisiri

TITLE OF INVENTION: Insecticidal Protein Toxins From Photorhabdus

NUMBER OF SEQUENCES: 88

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dow Agrosciences Patent Department

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: IN

COUNTRY: US

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/262,794A

FILING DATE: 02-OCT-2002

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,567

FILING DATE: 05-MAY-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/063,615

FILING DATE: 18-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/395,497

FILING DATE: 28-FEB-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/007,255

FILING DATE: 06-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/608,423

FILING DATE: 28-FEB-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/705,484

FILING DATE: 28-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296.93804

TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-251-5000

TELEFAX: 608-251-9166

INFORMATION FOR SEQ ID NO: 70:

SEQUENCE CHARACTERISTICS:

LENGTH: 22 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-10-262-794A-70

Query Match 28.9%; Score 28; DB 14; Length 22;

Best Local Similarity 40.0%; Pred. No. 1.3e+03;

Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

2 PNNLSKIAPKIVSQ 16

|||::|||

Query Match 28.9%; Score 28; DB 9; Length 23;
 Best Local Similarity 42.9%; Pred. No. 1.4e+03;
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEP 18
 DB 4 LKSFILSIQISKQEP 17

RESULT 154

US-10-097-065-263
 ; Sequence 263, Application US/10097065
 ; Publication No. US20030055236A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moore, Paul A. et al.
 ; TITLE OF INVENTION: 110 Human Secreted Proteins
 ; FILE REFERENCE: PZ021P1
 ; CURRENT APPLICATION NUMBER: US/10/097,065
 ; PRIOR FILING DATE: 2002-03-14
 ; PRIOR APPLICATION NUMBER: PCT/US98/27059
 ; PRIOR FILING DATE: 1998-12-17
 ; PRIOR APPLICATION NUMBER: 60/070,923
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,007
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,057
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,006
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,369
 ; PRIOR FILING DATE: 1997-12-19
 ; PRIOR APPLICATION NUMBER: 60/068,367
 ; PRIOR FILING DATE: 1997-12-19
 ; PRIOR APPLICATION NUMBER: 60/068,368
 ; PRIOR FILING DATE: 1997-12-19
 ; PRIOR APPLICATION NUMBER: 60/068,169
 ; PRIOR FILING DATE: 1997-12-19
 ; PRIOR APPLICATION NUMBER: 60/068,053
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,064
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,054
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,008
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,365
 ; PRIOR FILING DATE: 1997-12-19
 ; NUMBER OF SEQ ID NOS: 672
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 263
 ; LENGTH: 23
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-097-065-263

Query Match 28.9%; Score 28; DB 14; Length 23;
 Best Local Similarity 27.8%; Pred. No. 1.4e+03;
 Matches 5; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEP 18
 DB 5 EPOCGASRLSKWMLNISP 22

RESULT 155

US-10-372-876-263
 ; Sequence 263, Application US/10372876
 ; Publication No. US20030204071A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moore, Paul A. et al.
 ; TITLE OF INVENTION: 110 Human Secreted Proteins
 ; FILE REFERENCE: PZ021P1

; CURRENT APPLICATION NUMBER: US/10/372,876
 ; CURRENT FILING DATE: 2003-02-26
 ; PRIOR APPLICATION NUMBER: 09/334,595
 ; PRIOR FILING DATE: 1999-06-17
 ; PRIOR APPLICATION NUMBER: PCT/US98/27059
 ; PRIOR FILING DATE: 1998-12-17
 ; PRIOR APPLICATION NUMBER: 60/070,923
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,007
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,057
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,006
 ; PRIOR FILING DATE: 1997-12-18
 ; PRIOR APPLICATION NUMBER: 60/068,369
 ; PRIOR FILING DATE: 1997-12-19
 ; PRIOR APPLICATION NUMBER: 60/068,367
 ; PRIOR FILING DATE: 1997-12-19
 ; PRIOR APPLICATION NUMBER: 60/068,368
 ; PRIOR FILING DATE: 1997-12-19
 ; PRIOR APPLICATION NUMBER: 60/068,169
 ; PRIOR FILING DATE: 1997-12-19
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 672
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 263
 ; LENGTH: 23
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-372-876-263

Query Match 28.9%; Score 28; DB 14; Length 23;
 Best Local Similarity 27.8%; Pred. No. 1.4e+03;
 Matches 5; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEP 18
 DB 5 EPOCGASRLSKWMLNISP 22

RESULT 156

US-09-864-761-49103
 ; Sequence 49103, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aecomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 49103
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007539.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: SWISSPROT HIT: P52067, EVALUAE 1.70e+00
US-09-864-761-49103

Query Match      28.9%; Score 28; DB 9; Length 24;
Best Local Similarity 41.7%; Pred. No. 1.5e+03;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy      5 LNSKIAFKIVSQ 16
Db      6 LSARLAFVIFQ 17
      |:::| |::|
      |:::| |::|

RESULT 157
US-10-106-698-7023
; Sequence 7023, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7023
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (24)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-10-106-698-7023

Query Match      28.9%; Score 28; DB 14; Length 24;

Best Local Similarity 54.5%; Pred. No. 1.5e+03;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      3 NNLNSKIAFKI 13
Db      1 HELPSKISFEI 11
      :| | | | | | |
      :| | | | | | |

RESULT 158
US-10-424-599-167867
; Sequence 167867, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 167867
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_1225C.1.pap
US-10-424-599-167867

Query Match      28.9%; Score 28; DB 15; Length 24;
Best Local Similarity 33.3%; Pred. No. 1.5e+03;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy      5 LNSKIAFKIVSQ 16
Db      1 VNSYVVFMLNQ 12
      :| | | | | |
      :| | | | | |

RESULT 159
US-09-864-761-46865
; Sequence 46865, Application US/099664761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46865
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL023807.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
US-09-864-761-46865

```

```

Query Match      28.9%; Score 28; DB 9; Length 25;
Best Local Similarity 44.4%; Pred. No. 1.5e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      10 AFKIVSQEP 18
       |||:::
Db      3 AFKVMKEP 11

```

RESULT 160

```

US-09-752-723-12
; Sequence 12, Application US/09752723
; Publication No. US20020016976A1
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/752,723
; CURRENT FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
; OTHER INFORMATION: ATP-binding domain, V323A,I338A
US-09-752-723-12

```

```

Query Match      28.9%; Score 28; DB 9; Length 26;
Best Local Similarity 40.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

```

```

Qy      6 NSKIA--FKIVSQEP 18

```

```

Db      2 HEKLAQLYAVVSEEP 16
       :|:|:|:|:|:|:|

```

RESULT 161

```

US-09-752-723-13
; Sequence 13, Application US/09752723
; Publication No. US20020016976A1
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/752,723
; CURRENT FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
; OTHER INFORMATION: ATP-binding domain, V323A
US-09-752-723-13

```

```

Query Match      28.9%; Score 28; DB 9; Length 26;
Best Local Similarity 40.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

```

```

Qy      6 NSKIA--FKIVSQEP 18
       :|:|:|:|:|:|:|
Db      2 HEKLAQLYAVVSEEP 16

```

RESULT 162

```

US-10-428-254A-39
; Sequence 39, Application US/10428254A
; Publication No. US20030232420A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Cantor, Charles
; APPLICANT: Kammerer, Stefan
; APPLICANT: Taylor, Susan
; APPLICANT: Hamuro, Lora
; APPLICANT: Cook, Charles
; APPLICANT: Olson, Gary
; APPLICANT: Self, Christopher
; TITLE OF INVENTION: KINASE ANCHOR PROTEIN MUTAINS, PEPTIDES THEREOF, AND RELATED MET
; FILE REFERENCE: 24736-2065
; CURRENT APPLICATION NUMBER: US/10/428,254A
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 60/377,852
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 60/453,408
; PRIOR FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 24

```

```
; OTHER INFORMATION: Xaa is Ile or Val
US-10-428-254A-39

Query Match      28.9%; Score 28; DB 14; Length 27;
Best Local Similarity 72.7%; Pred. No. 1.7e+03;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVS 15
Db 12 LNWKIAKIVS 22

RESULT 163
US-09-864-761-41583
; Sequence 41583, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41583
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109976.21 SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
```

```
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EST_HUMAN HIT: AL049110.1, EVALUE 4.00e-09
US-09-864-761-41583

Query Match      28.9%; Score 28; DB 9; Length 29;
Best Local Similarity 33.3%; Pred. No. 1.8e+03;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 6 NSKIAFKIVSQE 17
Db 10 NSGISLKVQIED 21

RESULT 164
US-10-437-963-145993
; Sequence 145993, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 145993
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MET4530_46660C.1.pep
US-10-437-963-145993

Query Match      28.9%; Score 28; DB 16; Length 29;
Best Local Similarity 42.9%; Pred. No. 1.8e+03;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 4 HUNSKIAFKIVSQE 17
Db 12 HLYATIVFIIVTMQ 25

RESULT 165
US-09-900-147-6
; Sequence 6, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thanque, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

FEATURE:
US-09-900-147-6
OTHER INFORMATION: Description of Artificial Sequence: Synthetic

Query Match 28.9%; Score 28; DB 9; Length 30;
Best Local Similarity 46.2%; Pred. No. 1.9e+03;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSOE 17
|||:|:|:
Db 4 LNVLMANNIISKE 16

RESULT 166

US-10-437-963-151427
; Sequence 151427, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Audrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 151427
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_51571C.1.pap
US-10-437-963-151427

Query Match 28.9%; Score 28; DB 16; Length 32;
Best Local Similarity 71.4%; Pred. No. 2e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PNHLSK 8
|||:|:|:
Db 25 PNHQLK 31

RESULT 167

US-09-971-980-49
; Sequence 49, Application US/09971980
; Patent No. US20020164349A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B.
; APPLICANT: Yang, Joo-Sung
; TITLE OF INVENTION: Compositions and Methods of Using Capsid Protein From Flaviviruses
; TITLE OF INVENTION: Pestiviruses
; FILE REFERENCE: Upn-4105
; CURRENT APPLICATION NUMBER: US/09/971,980
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/237,885
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020164349A1e1 Sequence
US-09-971-980-49

Query Match 28.9%; Score 28; DB 9; Length 34;
Best Local Similarity 55.6%; Pred. No. 2.2e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PNHLSKIA 10
|||:|:|:
Db 14 PHLNLCVIS 22

RESULT 168

US-10-424-599-227024
; Sequence 227024, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 227024
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_47032C.1.pap
US-10-424-599-227024

Query Match 28.9%; Score 28; DB 15; Length 34;
Best Local Similarity 38.5%; Pred. No. 2.2e+03;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSOE 17
|||:|:|:
Db 4 LNNERSIKLVTK 16

RESULT 169

US-10-424-599-238675
; Sequence 238675, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 238675
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57549C.1.pap
US-10-424-599-238675

Query Match 28.9%; Score 28; DB 15; Length 35;
Best Local Similarity 36.8%; Pred. No. 2.2e+03;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSOEPA 19
|||:|:|:
Db 5 EPNFFSADKACNFFSDEQA 23


```
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34918
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL020996.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BP474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.7
; OTHER INFORMATION: EST HUMAN HIT: AA437353.1, EVALUE 3.00e-15
; OTHER INFORMATION: SWISSPROT HIT: P39098, EVALUE 9.00e-10
;
US-09-864-761-34918
```

```
Query Match      28.9%; Score 28; DB 9; Length 38;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      2 PNHLSNKIAP 11
        ||| | : |
Db      23 PNHDKKQSF 32
```

RESULT 175

```
US-10-144-929-247
; Sequence 247, Application US/10144929
; Publication No. US20030069405A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2014P1
; CURRENT APPLICATION NUMBER: US/10/144,929
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/251,329
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: PCT/US98/17044
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 247
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-929-247
```

```
Query Match      28.9%; Score 28; DB 14; Length 38;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      10 AFKIVSQEPA 19
        |:::| | | |
Db      12 AWRFLSQHPA 21
```

RESULT 176

```
US-10-144-929-247
; Sequence 247, Application US/10144929
; Publication No. US20040014954A9
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2014P1
; CURRENT APPLICATION NUMBER: US/10/144,929
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/251,329
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: PCT/US98/17044
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 247
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-929-247
```

```
Query Match      28.9%; Score 28; DB 15; Length 38;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      10 AFKIVSQEPA 19
        |:::| | | |
Db      12 AWRFLSQHPA 21
```

RESULT 177

```
US-09-962-756-203
; Sequence 203, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
```

```
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAEFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
```

```

; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 203
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-471-203

Query Match      28.9%; Score 28; DB 14; Length 39;
Best Local Similarity 50.0%; Pred. No. 2.5e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      10 AFKIVSQEPA 19
Db      14 AGKVASEQPA 23
      |||:|:|

RESULT 180
US-10-253-471-220
; Sequence 220, Application US/10253471
; Publication No. US20030236190A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 220
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-471-220

Query Match      28.9%; Score 28; DB 14; Length 39;
Best Local Similarity 60.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      10 AFKIVSQEPA 19
Db      14 AGKVASQPPA 23
      |||:|:|

RESULT 181
US-10-253-493-203
; Sequence 203, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:

```

```
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 203
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-493-203
```

```
Query Match 28.9%; Score 28; DB 15; Length 39;
Best Local Similarity 50.0%; Pred. No. 2.5e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 10 AFKIVSQEPA 19
|::|:|:|
Db 14 AGVASEQPA 23
```

```
RESULT 182
US-10-253-493-220
; Sequence 220, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 220
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-493-220
```

```
Query Match 28.9%; Score 28; DB 15; Length 39;
Best Local Similarity 60.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 10 AFKIVSQEPA 19
|::|:|:|
Db 14 AGVASEQPA 23
```

```
RESULT 183
US-09-942-024-62
; Sequence 62, Application US/09942024
; Publication No. US20030143650A1
; GENERAL INFORMATION:
; APPLICANT: Steward, Lance E.
```

```
; APPLICANT: Fernandez-Salas, Ester
; APPLICANT: Aoki, Kei Roger
; TITLE OF INVENTION: FRET Protease Assays For Botulinum
; FILE REFERENCE: Serotype A/E Toxins
; CURRENT APPLICATION NUMBER: US/09/942,024
; CURRENT FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-942-024-62
```

```
Query Match 28.9%; Score 28; DB 10; Length 40;
Best Local Similarity 35.3%; Pred. No. 2.6e+03;
Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
Qy 3 NHLNSKIAFKIVSQEPA 19
|::|:|:|
Db 19 NLVDSSVTFKTSRNLA 35
```

```
RESULT 184
US-09-942-098-62
; Sequence 62, Application US/09942098
; Publication No. US20030143651A1
; GENERAL INFORMATION:
; APPLICANT: Steward, Lance E.
; APPLICANT: Fernandez-Salas, Ester
; TITLE OF INVENTION: FRET Protease Assays For Clostridial
; FILE REFERENCE: P-AR 4802
; CURRENT APPLICATION NUMBER: US/09/942,098
; CURRENT FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-942-098-62
```

```
Query Match 28.9%; Score 28; DB 10; Length 40;
Best Local Similarity 35.3%; Pred. No. 2.6e+03;
Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
Qy 3 NHLNSKIAFKIVSQEPA 19
|::|:|:|
Db 19 NLVDSSVTFKTSRNLA 35
```

```
RESULT 185
US-10-261-161-87
; Sequence 87, Application US/10261161
; Publication No. US20040072270A1
; GENERAL INFORMATION:
; APPLICANT: Fernandez-Salas, Ester
; APPLICANT: Steward, Lance E.
; APPLICANT: Aoki, Kei Roger
; TITLE OF INVENTION: Cell-Based Fluorescence Resonance Energy
; FILE REFERENCE: P-AR 4804
; CURRENT APPLICATION NUMBER: US/10/261,161
; CURRENT FILING DATE: 2002-09-27
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Rattus sp.
```



```
RESULT 190
US-10-437-963-203041
; Sequence 203041, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203041
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_98262C.1.pep
US-10-437-963-203041

Query Match      28.9%; Score 28; DB 16; Length 41;
Best Local Similarity 83.3%; Pred. No. 2.7e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      3 NHLNSK 8
Db      10 NHLHSK 15

RESULT 191
US-09-933-767-557
; Sequence 557, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: PCT/US01/05614
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,836
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/193,170
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/205,258
; PRIOR FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: PCT/US98/11422
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/048,885
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,375
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,881
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,880
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,896
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,020
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,876
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,895
; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/048,884
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,894
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,971
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,882
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,899
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,893
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,900
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,901
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,892
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,915
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,019
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,373
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,875
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,374
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,917
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,949
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,883
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,897
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,898
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,962
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,963
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,877
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,878
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/085,925
```

```
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 557
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-933-767-557

Query Match      28.9%; Score 28; DB 10; Length 42;
Best Local Similarity 41.2%; Pred. No. 2.7e+03;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      3 NNLNSKIAPKIVSQRPA 19
Db      14 NYDQKNIFQCVSVRPA 30

RESULT 192
US-10-004-860-557
; Sequence 557, Application US/10004860
; Publication No. US20030065160A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/10/004,860
; CURRENT FILING DATE: 2001-12-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 557
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-860-557

Query Match      28.9%; Score 28; DB 14; Length 42;
Best Local Similarity 41.2%; Pred. No. 2.7e+03;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      3 NNLNSKIAPKIVSQRPA 19
Db      14 NYDQKNIFQCVSVRPA 30

RESULT 193
US-10-023-282-557
; Sequence 557, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
```


Query Match 28.9%; Score 28; DB 10; Length 45;
Best Local Similarity 46.2%; Pred. No. 3e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQEP 18
DB 27 NPKLFLNFSKEP 39

RESULT 198
US-10-378-029-86
; Sequence 86, Application US/10378029
; Publication No. US20040014087A1
; GENERAL INFORMATION:
; APPLICANT: HODGSON, David M.; LINCOLN, Stephen E.
; APPLICANT: RUSSO, Frank D.; SPIRO, Peter A.
; APPLICANT: BANVILLE, Steve C.; BRATCHER, Shawn R.
; APPLICANT: DUFOUR, Gerard E.; COHEN, Howard J.
; APPLICANT: ROSEN, Bruce; CHALUP, Michael S.
; APPLICANT: JACKSON, Jennifer L.; JONES, Annisa L.
; APPLICANT: YU, Jimmy Y.; GREENAWALT, Lila B.
; APPLICANT: PANZER, Scott R.; ROSEBERRY LINCOLN, Ann M.
; APPLICANT: WRIGHT, Rachel J.; DANIELS, Susan E.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1022-1 CIP
; CURRENT APPLICATION NUMBER: US/10/378,029
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 09/980,285
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/US00/15404
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: US 60/147,500
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,542
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,541
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,824
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,547
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,530
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,536
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,520
; PRIOR FILING DATE: 1999-08-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PERL Program
; SEQ ID NO 86
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040014087A1 430569.2.j.orf2
US-10-378-029-86

Query Match 28.9%; Score 28; DB 15; Length 45;
Best Local Similarity 25.0%; Pred. No. 3e+03;
Matches 4; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQEP 18
DB 18 HHLSVEILLERRIQSP 33

RESULT 199
US-10-424-599-175974
; Sequence 175974, Application US/10424599
; Publication No. US20040031072A1

GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175974
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129923C.1.pap
US-10-424-599-175974

Query Match 28.9%; Score 28; DB 15; Length 45;
Best Local Similarity 46.2%; Pred. No. 3e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQEP 18
DB 15 NSCLIFKVNLSLP 27

RESULT 200
US-10-424-599-200307
; Sequence 200307, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200307
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22903C.1.pap
US-10-424-599-200307

Query Match 28.9%; Score 28; DB 15; Length 45;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPHNLNSKIAFKIV 14
DB 21 EKNSLSSCIVFSAV 34

RESULT 201
US-10-424-599-213833
; Sequence 213833, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

Query Match 28.9%; Score 28; DB 10; Length 45;
Best Local Similarity 46.2%; Pred. No. 3e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQEP 18
DB 27 NPKLFLNFSKEP 39

RESULT 198
US-10-378-029-86
; Sequence 86, Application US/10378029
; Publication No. US20040014087A1
; GENERAL INFORMATION:
; APPLICANT: HODGSON, David M.; LINCOLN, Stephen E.
; APPLICANT: RUSSO, Frank D.; SPIRO, Peter A.
; APPLICANT: BANVILLE, Steve C.; BRATCHER, Shawn R.
; APPLICANT: DUFOUR, Gerard E.; COHEN, Howard J.
; APPLICANT: ROSEN, Bruce; CHALUP, Michael S.
; APPLICANT: JACKSON, Jennifer L.; JONES, Annisa L.
; APPLICANT: YU, Jimmy Y.; GREENAWALT, Lila B.
; APPLICANT: PANZER, Scott R.; ROSEBERRY LINCOLN, Ann M.
; APPLICANT: WRIGHT, Rachel J.; DANIELS, Susan E.
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1022-1 CIP
; CURRENT APPLICATION NUMBER: US/10/378,029
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 09/980,285
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/US00/15404
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: US 60/147,500
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,542
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,541
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,824
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,547
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,530
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,536
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 60/147,520
; PRIOR FILING DATE: 1999-08-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PERL Program
; SEQ ID NO 86
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040014087A1 430569.2.j.orf2
US-10-378-029-86

Query Match 28.9%; Score 28; DB 15; Length 45;
Best Local Similarity 25.0%; Pred. No. 3e+03;
Matches 4; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQEP 18
DB 18 HHLSVEILLERRIQSP 33

RESULT 199
US-10-424-599-175974
; Sequence 175974, Application US/10424599
; Publication No. US20040031072A1

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 213833
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(45)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_35117C.1.pap
US-10-424-599-213833

Query Match 28.9%; Score 28; DB 15; Length 45;
Best Local Similarity 33.3%; Pred. No. 3e+03;
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 PNHLNSKIAPKI 13
Db 34 PINSINPEKFOV 45

RESULT 202
US-10-424-599-227298
; Sequence 227298, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 227298
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_4727C.1.pap
US-10-424-599-227298

Query Match 28.9%; Score 28; DB 15; Length 45;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PNHLNSKI 9
Db 18 PNHKNSL 25

RESULT 203
US-10-424-599-277857
; Sequence 277857, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 277857
; LENGTH: 45
; TYPE: PRT

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_92928C.1.pap
US-10-424-599-277857

Query Match 28.9%; Score 28; DB 15; Length 45;
Best Local Similarity 30.8%; Pred. No. 3e+03;
Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 7 SKIAFKIVSQEPA 19
Db 28 NKIVHEVITSNPA 40

RESULT 204
US-10-004-378A-59
; Sequence 59, Application US/10004378A
; Publication No. US20030228301A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazarzyna
; APPLICANT: Perna, Amanda
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corrine A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Bsha A
; TITLE OF INVENTION: No. US20030228301A1el Human Proteins, Polynucleotides Encoding T
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004,378A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 59
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-378A-59

```

Query Match 28.9%; Score 28; DB 14; Length 46;
Best Local Similarity 31.2%; Pred. NO. 3e+03;
Matches 5: Conservative 3; Mismatches 8; Indels

Qy 2 PNHLNSKIAFKIVSQE 17
||| : : :
Db 27 PEHLTDYFWSIPEEE 42

```

RESULT 205
US-10-424-599-221525
; Sequence 221525, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 221525
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_42066C.1.pep
US-10-424-599-221525

```

Query Match 28.9%; Score 28; DB 15; Length 46;
Best Local Similarity 60.0%; Pred. No. 3e+03;
Matches 6: Conservative 0: Mismatches 4: Indels

Qy	4	HLNSKIAFKI	13
Db	34	HYVSKCCEKI	43

```

RESULT 206
US-10-424-599-254250
; Sequence 254250, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254250
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(46)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71611C.1.pep
US-10-424-599-254250

```

Query Match 28.9%; Score 28; DB 15; Length 46;
Best Local Similarity 43.8%; Pred. No. 3e+03;
Matches 7: Conservative 2; Mismatches 5; Indels

Qy 1 EPNHLSK--IAFKIV 14
||| : |||
Db 13 EPHYLEERDVVGEKHV 28

```

RESULT 207
US-10-424-599-272996
; Sequence 272996, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 272996
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89538C.1.pep
US-10-424-599-272996

```

Query Match	28.9%	Score 28;	DB 15;	Length 46;
Best Local Similarity	23.1%	Pred. No. 3e+03;		
Matches	3:	Conservative	7: Mismatches	3: Indels
				Gaps 0;

Qy 3 NHLNSKIAFKIVS 15
:|: ::|:::
Db 27 DHL\$KDVSLKLI 39

RESULT 208
US-09-864-761-33481
; Sequence 33481, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rankel, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669

```

RESULT 209
US-10-106-698-5380
; Sequence 5380, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5380
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5380

```

Qy 4 HLNSKIAFKIVSO 16

```
Db      14 HLLSQVQFMTYSQ 26
|||::| ||
RESULT 212
US-10-424-599-242667
; Sequence 242667, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 242667
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_61158C.1.pep
US-10-424-599-242667
Query Match      28.9%; Score 28; DB 15; Length 48;
Best Local Similarity 58.3%; Pred. No. 3.2e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY      4 HLNSKIAPKIVS 15
||| |||:|
Db      7 HLNRN--FKIIS 16

RESULT 213
US-10-437-963-113492
; Sequence 113492, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 113492
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_17274C.1.pep
US-10-437-963-113492
Query Match      28.9%; Score 28; DB 16; Length 48;
Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      4 HLNSKIAPKI 13
||| |||:|
Db      7 HLGSALNFWSI 16

RESULT 214
US-10-424-599-188532
; Sequence 188532, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 188532
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_141259C.1.pep
US-10-424-599-188532
Query Match      28.9%; Score 28; DB 15; Length 49;
Best Local Similarity 40.0%; Pred. No. 3.2e+03;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 PNHLSKIAF 11
|||:|:|
Db      5 PSHLHLSJSF 14

RESULT 215
US-10-424-599-217183
; Sequence 217183, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 217183
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_38143C.1.pep
US-10-424-599-217183
Query Match      28.9%; Score 28; DB 15; Length 49;
Best Local Similarity 46.2%; Pred. No. 3.2e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      2 PNHLSKIAFKIV 14
|||:|:|
Db      16 PNFLENTMNAAKAI 28

RESULT 216
US-10-424-599-224025
; Sequence 224025, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 224025
 ; LENGTH: 49
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_44323C.1.pep
 US-10-424-599-224025

Query Match 28.9%; Score 28; DB 15; Length 49;
 Best Local Similarity 33.3%; Pred. No. 3.2e+03;
 Matches 5; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 HLNSKIAFKIVSQEP 18
 ||||| : :
 Db 1 HLNSNLNYFFTKNPP 15

RESULT 217
 US-10-083-357-970
 ; Sequence 370, Application US/10083357
 ; Publication No. US20030054370A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Qiangdong Zeng et al.
 ; TITLE OF INVENTION: Systemic Discovery of New Genes
 ; FILE REFERENCE: 032796-090
 ; CURRENT APPLICATION NUMBER: US/10/083,357
 ; CURRENT FILING DATE: 2002-02-27
 ; NUMBER OF SEQ ID NOS: 1346
 ; SEQ ID NO 970
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-10-083-357-970

Query Match 28.9%; Score 28; DB 14; Length 50;
 Best Local Similarity 50.0%; Pred. No. 3.3e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSK 8
 ||||| : :
 Db 23 EPNHCKNR 30

RESULT 218
 US-10-424-599-188812
 ; Sequence 188812, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 188812
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_141511C.1.pep
 US-10-424-599-188812

Query Match 28.9%; Score 28; DB 15; Length 50;
 Best Local Similarity 41.7%; Pred. No. 3.3e+03;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 EPNHLSKIAFK 12
 : ||||| : :
 Db 33 DPKLNAALSLK 44

RESULT 219
 US-10-424-599-208956
 ; Sequence 208956, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 208956
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_30715C.1.pep
 US-10-424-599-208956

Query Match 28.9%; Score 28; DB 15; Length 50;
 Best Local Similarity 44.4%; Pred. No. 3.3e+03;
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 NHLNSKIAF 11
 ||||| : :
 Db 33 NHLNSLTSF 41

RESULT 220
 US-10-437-963-128124
 ; Sequence 128124, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 128124
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_30508C.1.pep
 US-10-437-963-128124

Query Match 28.9%; Score 28; DB 16; Length 50;
 Best Local Similarity 38.5%; Pred. No. 3.3e+03;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 SKIAFKIVSQEPA 19
 ||||| : :
 Db 35 SEIEFEFYAQNPS 47

RESULT 221
US-10-437-963-160977
; Sequence 160977, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 160977
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_60204C.1.pep
US-10-437-963-160977

Query Match 28.9%; Score 28; DB 16; Length 50;
Best Local Similarity 35.3%; Pred. No. 3.3e+03;
Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQE 17
DB 14 ETHTSSHIFQSKSISSD 30

RESULT 222
US-10-437-963-194803
; Sequence 194803, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 194803
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90813C.1.pep
US-10-437-963-194803

Query Match 28.9%; Score 28; DB 16; Length 50;
Best Local Similarity 54.5%; Pred. No. 3.3e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 HLNSKIAPKIV 14
DB 34 HLTSDLAFDSV 44

RESULT 223
US-10-424-599-209626
; Sequence 209626, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 209626
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_3131C.1.pep
US-10-424-599-209626

Query Match 28.4%; Score 27.5; DB 15; Length 41;
Best Local Similarity 43.8%; Pred. No. 3.2e+03;
Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 1 EPNHLSKIAPKIVSQ 16
DB 4 KPHHANT-ISFVIHKQ 18

RESULT 224
US-10-424-599-249169
; Sequence 249169, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249169
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6702C.1.pep
US-10-424-599-249169

Query Match 28.4%; Score 27.5; DB 15; Length 41;
Best Local Similarity 36.8%; Pred. No. 3.2e+03;
Matches 7; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1 EPNHLSKIAPKIVSQEP 18
DB 6 EPRHLPFSRGRHVLSP 24

RESULT 225
US-10-437-963-110359
; Sequence 110359, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 110359
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_1442C.1.pep
US-10-437-963-110359

Query Match      28.4%; Score 27.5; DB 16; Length 44;
Best Local Similarity 35.0%; Pred. No. 3.5e+03;
Matches 7; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

Qy 1 EPNHLSK-IAFKIVSQEPA 19
Db 24 EBNHLSKRGEDISVLQGEPA 43

RESULT 226
US-10-424-599-163952
; Sequence 163952, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163952
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_119067C.1.pep
US-10-424-599-163952

Query Match      28.4%; Score 27.5; DB 15; Length 45;
Best Local Similarity 27.8%; Pred. No. 3.6e+03;
Matches 5; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Qy 2 PNHLSKIAFKIVSQEPA 19
Db 4 PHHQKQVA-KTMTKNPS 20

RESULT 227
US-10-424-599-206168
; Sequence 206168, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
```

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206168
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28197C.1.pep
US-10-424-599-206168

Query Match      28.4%; Score 27.5; DB 15; Length 45;
Best Local Similarity 41.2%; Pred. No. 3.6e+03;
Matches 7; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

Qy 2 PNH---LNSKIAPKIVS 15
Db 2 PNHREMLTSRVRRSIVN 18

RESULT 228
US-10-437-963-147299
; Sequence 147299, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147299
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4783C.1.pep
US-10-437-963-147299

Query Match      28.4%; Score 27.5; DB 16; Length 45;
Best Local Similarity 60.0%; Pred. No. 3.6e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 3 NHL-NSKIAP 11
Db 4 NHLHNKKVGF 13

RESULT 229
US-10-424-599-232526
; Sequence 232526, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 232526
; LENGTH: 50
; TYPE: PRT
```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51996C.1.pep
US-10-424-599-232526

Query Match 28.4%; Score 27.5; DB 15; Length 50;
Best Local Similarity 38.5%; Pred. No. 4e+03; 1; Indels 3; Gaps 1;
Matches 5; Conservative 4; Mismatches 3; Indels 1;

Qy 2 PNHLNSKIAFKIV 14
Db 12 PNHL--LGFLQI 21
||||| :|::|

RESULT 230

US-09-946-175-7
; Sequence 7, Application US/09946175
; Patent No. US20020106671A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert
; TITLE OF INVENTION: IDENTIFICATION OF A CAPACITATIVE CALCIUM
; CHANNEL IN ANTIGEN PRESENTING CELLS AND USES THEREOF
; FILE REFERENCE: M1sh1(310800)
; CURRENT APPLICATION NUMBER: US/09/946,175
; CURRENT FILING DATE: 2001-09-05
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-175-7

Query Match 27.8%; Score 27; DB 9; Length 21;
Best Local Similarity 47.1%; Pred. No. 1.8e+03;
Matches 8; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

Qy 1 EPNHLNSKIAFKIVSQE 17
Db 5 EPGKLSQKI--KVVQLQE 19
||| :||| :|

RESULT 231

US-09-883-343A-12
; Sequence 12, Application US/09883343A
; Publication No. US20030039632A1
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Rodney W.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aislin
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: NO. US20030039632A1elBacteriocins, Transport and Vector System an
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/09/883,343A
; CURRENT FILING DATE: 2001-06-19
; PRIOR FILING DATE: 2001-06-19
; PRIOR FILING DATE: 1997-09-05
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Lactococcin A;
US-09-883-343A-12

Query Match 27.8%; Score 27; DB 10; Length 21;
Best Local Similarity 38.5%; Pred. No. 1.8e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQE 17
Db 1 MKNQLNFNIVSDE 13
:|::| :|::|

RESULT 232

US-09-883-343A-80
; Sequence 80, Application US/09883343A
; Publication No. US20030039632A1
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aislin
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: NO. US20030039632A1elBacteriocins, Transport and Vector System a
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/09/883,343A
; CURRENT FILING DATE: 2001-06-19
; PRIOR FILING DATE: 1997-09-05
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Lactococcin M
US-09-883-343A-80

Query Match 27.8%; Score 27; DB 10; Length 21;
Best Local Similarity 30.8%; Pred. No. 1.8e+03;
Matches 4; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQE 17
Db 1 MKNQLNFELSEE 13
:|::| :|::|

RESULT 233

US-10-097-079-56
; Sequence 56, Application US/10097079
; Publication No. US20020132973A1
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/097,079
; APPLICATION NUMBER: 13-Mar-2002
; FILING DATE: 13-Mar-2002


```

; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/228,990
; FILING DATE: <Unknown>
; NAME: OBLON, NORMAN F.
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: NO. US20020132973A1 Relevant
; MOLECULE TYPE: Peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 9..13
; OTHER INFORMATION: /product= "OTHER"
; /note= "The side chains of Lys at position 9 and Asp at position
; 13 are linked by an amide bond."
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 22
; OTHER INFORMATION: /product= "OTHER"
; /note= "This C-terminal amino acid is an amide, i.e., CONH2."
; US-10-097-079-56
;
; Query Match 27.8%; Score 27; DB 13; Length 22;
; Best Local Similarity 100.0%; Pred. No. 1.9e+03;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 4 HLNSK 8
; DB 5 HLNSK 9
;
; RESULT 234
; US-09-925-442-26
; Sequence 26, Application US/09925442
; Patent No. US20020103346A1
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; KOCK, MICHAEL
; BREDEHORST, REINHORST
; FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/925,442
; FILING DATE: 10-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/017,947
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

```

; APPLICATION NUMBER: 09/017,947
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
; US-09-925-442-26
;
; Query Match 27.8%; Score 27; DB 9; Length 23;
; Best Local Similarity 38.9%; Pred. No. 2e+03;
; Matches 7; Conservative 3; Mismatches 6; Indels 2; Gaps 1;
;
; QY 2 PNH--LNSKIAFKIVSQE 17
; DB 5 PTHQDLNLDISIKLPERE 22
;
; RESULT 235
; US-09-925-442-28
; Sequence 28, Application US/09925442
; Patent No. US20020103346A1
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; BREDEHORST, REINHORST
; KOCK, MICHAEL
; FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/925,442
; FILING DATE: 10-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/017,947
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-925-442-28
Query Match      27.8%; Score 27; DB 9; Length 23;
Best Local Similarity 35.3%; Pred. No. 2e+03;
Matches 6; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

QY      2 PNH--LNSKIAFKIVSO 16
Db      5 PDKDLNMDVSFHLPSR 21

RESULT 236
US-10-097-079-57
; Sequence 57, Application US/10097079
; Publication No. US20020132973A1
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/097,079
; FILING DATE: 13-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/228,990
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: No. US20020132973A1 Relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 23
; OTHER INFORMATION: /product= "OTHER"
; /note= "This C-terminal amino acid is an amide, i.e., CONH2."
; SEQUENCE DESCRIPTION: SEQ ID NO: 57:
US-10-097-079-57
Query Match      27.8%; Score 27; DB 13; Length 23;
Best Local Similarity 100.0%; Pred. No. 2e+03;

; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-994-595-111
; Sequence 111, Application US/09994595
; Publication No. US20030039981A1
; GENERAL INFORMATION:
; APPLICANT: Bhattacherjee, J.
; APPLICANT: Bhattacherjee, Vaquer
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DETECTING FUNGAL PATHOGENS IN
; TITLE OF INVENTION: A BIOLOGICAL SAMPLE
; FILE REFERENCE: 96,247-A
; CURRENT APPLICATION NUMBER: US/09/994,595
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 08/650,809
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 111
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Polypeptide segment of GRGB_IBACB shown in Figure 4.
US-09-994-595-111
Query Match      27.8%; Score 27; DB 10; Length 24;
Best Local Similarity 62.5%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4 HLNSKIAF 11
Db      9 HLKDKFAF 16

RESULT 238
US-10-097-079-58
; Sequence 58, Application US/10097079
; Publication No. US20020132973A1
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/097,079
; FILING DATE: 13-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/228,990
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: No. US20020132973A1 Relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 23
; OTHER INFORMATION: /product= "OTHER"
; /note= "This C-terminal amino acid is an amide, i.e., CONH2."
; SEQUENCE DESCRIPTION: SEQ ID NO: 57:
US-10-097-079-57
Query Match      27.8%; Score 27; DB 13; Length 23;
Best Local Similarity 100.0%; Pred. No. 2e+03;
```

```
;
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: No. US20020132973A1 Relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 11,15
; OTHER INFORMATION: /product= "OTHER"
; /note= "The side chains of Lys at position 11 and Asp at
; position 15 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 24
; OTHER INFORMATION: /product= "OTHER"
; /note= "This C-terminal amino acid is an amide, i.e., CONH2."
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-10-097-079-58
Query Match 27.8%; Score 27; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 2.1e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 HLNSK 8
Db 7 HLNSK 11
RESULT 239
US-10-097-079-59
; Sequence 59, Application US/10097079
; Publication No. US20020132973A1
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/097,079
; FILING DATE: 13-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/228,990
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
```

```
;
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: No. US20020132973A1 Relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 2
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 12,16
; OTHER INFORMATION: /product= "OTHER"
; /note= "The side chains of Lys at position 12 and Asp at
; position 16 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 25
; OTHER INFORMATION: /product= "OTHER"
; /note= "This C-terminal amino acid is an amide, i.e., CONH2."
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-10-097-079-59
Query Match 27.8%; Score 27; DB 13; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.2e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 HLNSK 8
Db 8 HLNSK 12
RESULT 240
US-10-449-831A-45
; Sequence 45, Application US/10449831A
; Publication No. US20040029179A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; TITLE OF INVENTION: Higher molecular weight entities and uses therefor
; FILE REFERENCE: 2385978
; CURRENT APPLICATION NUMBER: US/10/449,831A
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: USN 60/384878
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 237
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
; LENGTH: 25
; TYPE: PPT
; ORGANISM: Haemophilus influenzae
US-10-449-831A-45
Query Match 27.8%; Score 27; DB 15; Length 25;
Best Local Similarity 50.0%; Pred. No. 2.2e+03; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 3 NHLNSKIAFKIVSQ 16
Db 9 NFLTACISLGIVSQ 22
RESULT 241
US-09-752-723-3
; Sequence 3, Application US/09752723
; Publication No. US20020016976A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/752,723
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Rous sarcoma virus
; FEATURE:
; OTHER INFORMATION: ATP binding domain of v-Src, residues 318-343
US-09-752-723-3

Query Match      27.8%; Score 27; DB 9; Length 26;
Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
Db 9 YAVVSEEP 16

RESULT 242
US-09-752-723-10
; Sequence 10, Application US/09752723
; Publication No. US20020016976A1
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/752,723
; CURRENT FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
; OTHER INFORMATION: ATP-binding domain, I338G
US-09-752-723-10

Query Match      27.8%; Score 27; DB 9; Length 26;
Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
Db 9 YAVVSEEP 16
```

RESULT 243

```
US-09-752-723-11
; Sequence 11, Application US/09752723
; Publication No. US20020016976A1
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/752,723
; CURRENT FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
; OTHER INFORMATION: ATP-binding domain, I338A
US-09-752-723-11

Query Match      27.8%; Score 27; DB 9; Length 26;
Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
Db 9 YAVVSEEP 16

RESULT 244
US-09-752-723-14
; Sequence 14, Application US/09752723
; Publication No. US20020016976A1
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/752,723
; CURRENT FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
; OTHER INFORMATION: ATP-binding domain, I338S
US-09-752-723-14

Query Match      27.8%; Score 27; DB 9; Length 26;
Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
```

```

Db          : :||:|
           9 YAVVSEEP 16

Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
   : :||:|
Db  9 YAVVSEEP 16

RESULT 245
US-09-752-723-19
; Sequence 19, Application US/09752723
; Publication No. US20020016976A1
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: Nucleotide Triphosphate Substrates
; CURRENT APPLICATION NUMBER: US/09/752,723
; CURRENT FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-03-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: ATP binding domain of Fyn, residues 319-344
US-09-752-723-19

Query Match 27.8%; Score 27; DB 9; Length 26;
Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
   : :||:|
Db  9 YAVVSEEP 16

RESULT 246
US-09-752-723-20
; Sequence 20, Application US/09752723
; Publication No. US20020016976A1
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: Nucleotide Triphosphate Substrates
; CURRENT APPLICATION NUMBER: US/09/752,723
; CURRENT FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-03-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant Fyn
US-09-752-723-20

Query Match 27.8%; Score 27; DB 9; Length 26;

Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
   : :||:|
Db  9 YAVVSEEP 16

RESULT 247
US-09-347-064-14
; Sequence 14, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; TITLE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acids
US-09-347-064-14

Query Match 27.8%; Score 27; DB 9; Length 26;
Best Local Similarity 38.5%; Pred. No. 2.3e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 EPNHLNSKIAPKI 13
   | :||:|
Db  3 ESTFKNTEISFKL 15

RESULT 248
US-09-985-157-3
; Sequence 3, Application US/09985157
; Patent No. US20020146797A1
; GENERAL INFORMATION:
; APPLICANT: SHOKAT, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases Which Can Utilize Modified Nucleotide
; FILE REFERENCE: 51538-5002-05
; CURRENT APPLICATION NUMBER: US/09/985,157
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: US 08/797,552
; PRIOR FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Rous sarcoma virus
US-09-985-157-3
```

Query Match 27.8%; Score 27; DB 9; Length 26;
Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
: : : : :
Db 9 YAVVSEEP 16

RESULT 249

US-10-097-079-60
; Sequence 60, Application US/10097079
; Publication No. US20020132973A1
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/097,079
; FILING DATE: 13-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/228,990
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: No. US20020132973A1 Relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 3
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 13..17
; OTHER INFORMATION: /product= "OTHER"
; /note= "The side chains of Lys at position 13 and Asp at
; position 17 are linked by an amide bond."
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 26
; OTHER INFORMATION: /product= "OTHER"
; /note= "This C-terminal amino acid is an amide, i.e., CONH2."
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-10-097-079-60

Query Match 27.8%; Score 27; DB 13; Length 26;

Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 HLNSK 8
: : : : :
Db 9 HLNSK 13

RESULT 250

US-10-044-967-3
; Sequence 3, Application US/10044967
; Publication No. US20030073218A1
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: High Affinity Kinase Inhibitors for Target Validation
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 51538-5001-US
; CURRENT APPLICATION NUMBER: US/10/044,967
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 09/480,993
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/115,340
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Rous sarcoma virus
US-10-044-967-3

Query Match 27.8%; Score 27; DB 14; Length 26;
Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
: : : : :
Db 9 YAVVSEEP 16

Search completed: October 19, 2004, 19:25:15
Job time : 141 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: October 19, 2004, 19:05:00 ; Search time 38 Seconds
(without alignments)
48.108 Million cell updates/sec

Title: US-10-799-005A-1
Perfect score: 97
Sequence: 1 EPNHLSKIAFKIVSQEPA 19
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 8622
Minimum DB seq length: 19
Maximum DB seq length: 50
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries
Database : PIR 79:*

1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	35.1	38	T14885	hypothetical prote
2	33	34.0	20	A53875	creatine kinase (G
3	33	34.0	50	D82660	hypothetical prote
4	30.5	31.4	44	B57473	caveolin beta form
5	30	30.9	47	JC6158	viral capsid prote
6	29	29.9	33	H82096	hypothetical prote
7	29	29.9	40	G82484	hypothetical prote
8	28.5	29.4	49	H87397	hypothetical prote
9	28	28.9	24	S47281	hypothetical prote
10	28	28.9	34	D95189	hypothetical prote
11	28	28.9	36	S17507	cytokine - rabbit
12	27	27.8	30	PQ0444	hypothetical prote
13	27	27.8	33	S22603	ribosomal protein
14	27	27.8	36	A61235	fibroblast-activat
15	27	27.8	38	G70164	hypothetical prote
16	27	27.8	43	S70031	placental lactogen
17	27	27.8	47	E72258	hypothetical prote
18	27	27.8	48	S62189	phosphatidylserine
19	26.5	27.3	34	S70246	nifW2 protein - An
20	26.5	27.3	37	D82636	hypothetical prote
21	26	26.8	27	S78355	hypothetical prote
22	26	26.8	31	E82446	hypothetical prote
23	26	26.8	34	JQ1620	hypothetical 4.3k
24	26	26.8	35	S82051	hypothetical prote
25	26	26.8	36	S21594	Ig heavy chain V r
26	26	26.8	38	A60867	properdin - rabbit
27	26	26.8	42	H70246	hypothetical prote
28	26	26.8	43	I36946	haptoglobin pseudo
29	26	26.8	43	E82428	hypothetical prote

30	26	26.8	44	2	A19434	hypothetical prote
31	26	26.8	45	2	S66648	major acute phase
32	26	26.8	48	2	AC1020	hypothetical prote
33	26	26.8	48	2	I38223	protein-serine/thr
34	26	26.8	50	2	E81927	hypothetical prote
35	25.5	26.3	42	1	WMBPPI	gene i protein - p
36	25	25.8	20	2	S08605	hypothetical prote
37	25	25.8	20	2	B46174	RNA-binding protei
38	25	25.8	21	2	JU0386	nitrile hydratase
39	25	25.8	26	2	B46451	RNA-directed RNA p
40	25	25.8	30	2	A22977	delta-endotoxin -
41	25	25.8	36	2	A69827	hypothetical prote
42	25	25.8	36	2	D82617	hypothetical prote
43	25	25.8	41	2	D69213	hypothetical prote
44	25	25.8	42	2	I48852	tropomyosin 5 - mo
45	25	25.8	44	2	D37286	olfactory receptor
46	25	25.8	44	2	B72266	hypothetical prote
47	25	25.8	45	2	T52125	R2R3-MYB transcrip
48	25	25.8	46	2	S78243	hypothetical prote
49	25	25.8	46	2	T05086	hypothetical prote
50	25	25.8	47	2	S31005	gene 60 protein -
51	25	25.8	48	2	A30504	Ig gamma heavy cha
52	25	25.8	49	2	S72793	cysteine synthase
53	25	25.8	50	2	H97798	hypothetical prote
54	24.5	25.3	32	2	F95089	hypothetical prote
55	24.5	25.3	34	2	A39965	macrocinn-O-methylc
56	24.5	25.3	38	2	T01741	hypothetical prote
57	24.5	25.3	42	2	JT0514	Ig heavy chain V-I
58	24.5	25.3	50	2	I51287	calcium-binding pr
59	24	24.7	25	2	I77474	myoglobin - Serico
60	24	24.7	26	4	A01640	hypothetical prote
61	24	24.7	28	2	S72460	ribosomal protein
62	24	24.7	29	1	GCCB	glucagon - Chinch
63	24	24.7	31	2	D95224	hypothetical prote
64	24	24.7	38	2	AD0740	hypothetical prote
65	24	24.7	40	2	S72464	ribosomal protein
66	24	24.7	41	2	S25728	beta-lactamase (EC
67	24	24.7	41	2	T18202	hypothetical prote
68	24	24.7	41	2	AC1594	hypothetical prote
69	24	24.7	41	2	A11645	protein gp22 (Bact
70	24	24.7	44	2	S72462	ribosomal protein
71	24	24.7	46	2	D49281	pol protein - simi
72	24	24.7	46	2	A49281	pol protein - simi
73	24	24.7	46	2	F49281	pol protein - simi
74	24	24.7	46	2	E49281	pol protein - simi
75	24	24.7	46	2	C49281	pol protein - simi
76	24	24.7	47	2	S14022	hypothetical prote
77	24	24.7	48	2	H64544	ribosomal protein
78	24	24.7	48	2	G71964	ribosomal protein
79	24	24.7	50	2	C72257	hypothetical prote
80	24	24.7	50	2	S06495	penicillin-binding
81	24	24.7	50	2	G82540	hypothetical prote
82	23.5	24.2	43	2	S21694	thymosin beta-12 -
83	23.5	24.2	48	2	S58644	tubulin beta chain
84	23.5	24.2	48	2	G86209	protein F22G5.12
85	23	23.7	23	2	S07700	T-cell receptor al
86	23	23.7	25	2	S39391	calpain II heavy c
87	23	23.7	29	2	I37301	MHC class II histo
88	23	23.7	30	2	S1617	ribosomal protein
89	23	23.7	30	2	G82528	hypothetical prote
90	23	23.7	31	2	E95866	probable pyrroloqu
91	23	23.7	33	2	C82287	hypothetical prote
92	23	23.7	35	2	F82443	hypothetical prote
93	23	23.7	36	2	H70251	hypothetical prote
94	23	23.7	37	2	S70801	probable membrane-
95	23	23.7	37	2	A23617	conglutinin dela-2
96	23	23.7	38	2	T06924	photosystem II pro
97	23	23.7	39	2	S72459	ribosomal protein
98	23	23.7	39	2	C97513	hypothetical prote
99	23	23.7	40	2	I49405	beta-galactoside-b
100	23	23.7	41	2	E83742	hypothetical prote
101	23	23.7	41	2	A97948	hypothetical prote
102	23	23.7	44	2	I37286	olfactory receptor

103	23	23.7	44	2	E37286	olfactory receptor	176	22	22.7	48	2	S61469	p83/100 protein -
104	23	23.7	44	2	E85938	hypothetical prote	177	22	22.7	48	2	S61472	p83/100 protein -
105	23	23.7	45	2	AG2286	cytochrome b559 be	178	22	22.7	49	2	B48396	ribosomal protein
106	23	23.7	45	2	A61295	creatine kinase (E	179	22	22.7	50	2	A64019	hypothetical prote
107	23	23.7	45	2	B61295	creatine kinase (E	180	22	22.7	50	2	H90596	hypothetical prote
108	23	23.7	45	2	H83816	hypothetical prote	181	21.5	22.2	19	2	A37968	neural surface pro
109	23	23.7	46	2	T07459	hypothetical prote	182	21.5	22.2	20	2	T48861	leader peptide [im
110	23	23.7	47	2	B36626	hypothetical prote	183	21.5	22.2	27	2	S28940	cyclic nucleotide-
111	23	23.7	47	2	T17087	osteogenic protein	184	21.5	22.2	28	2	S37683	protein IEF SSP 91
112	23	23.7	47	2	R81833	probable ts elemen	185	21.5	22.2	34	2	S12554	hydroxymethylgluta
113	23	23.7	48	2	A24941	hypothetical prote	186	21.5	22.2	35	2	PS0364	protein-tyrosine-p
114	23	23.7	48	2	F70247	storage protein -	187	21	21.6	20	2	A60365	tryptophan inhibito
115	23	23.7	48	2	R22224	hypothetical prote	188	21	21.6	20	2	S71593	serine proteinase
116	23	23.7	48	2	B69173	ribosomal protein	189	21	21.6	20	2	D1RT	dental fluid tra
117	23	23.7	49	2	G69538	conserved hypotnet	190	21	21.6	21	2	S47207	T-cell receptor J
118	23	23.7	50	2	A16662	probable pheromone	191	21	21.6	22	2	D30609	ig kappa chain V-I
119	23	23.7	50	2	F24609	hypothetical prote	192	21	21.6	22	2	B60691	phycobilisome 29K
120	23	23.7	50	2	D31753	hypothetical Oct-2	193	21	21.6	26	2	A37476	hypothetical prote
121	22.5	23.2	36	2	G82817	hypothetical prote	194	21	21.6	27	2	A45140	fatty-acyl-ethyl-e
122	22.5	23.2	48	2	F82294	hypothetical prote	195	21	21.6	28	2	C83969	hypothetical prote
123	22.5	23.2	50	2	S26130	outer membrane pro	196	21	21.6	28	2	F97000	hypothetical prote
124	22	22.7	20	2	S00774	kinase-related tra	197	21	21.6	29	2	A35891	carcinoembryonic a
125	22	22.7	20	2	C49164	chromogranin-B - r	198	21	21.6	30	2	A44913	34K core flagella
126	22	22.7	22	2	PT0070	hypothetical prote	199	21	21.6	30	2	B55020	hypothetical prote
127	22	22.7	24	1	S58242	pyrroloquinoline q	200	21	21.6	30	2	A12205	hypothetical prote
128	22	22.7	25	2	A32203	dihydrofolate redu	201	21	21.6	31	2	I54515	pre-B cell Ig lamb
129	22	22.7	26	2	I51597	albumin - African	202	21	21.6	31	2	S26232	ribosomal protein
130	22	22.7	30	2	S55462	mer5 protein homol	203	21	21.6	31	2	S31176	microtubule-associ
131	22	22.7	30	2	A05315	pancreatic ribonuc	204	21	21.6	31	2	A60122	fimbrial protein -
132	22	22.7	30	2	A47607	immunogenic protei	205	21	21.6	31	2	E95140	hypothetical prote
133	22	22.7	30	2	F82209	hypothetical prote	206	21	21.6	31	2	D70223	hypothetical prote
134	22	22.7	31	2	PH0248	T-cell receptor Vb	207	21	21.6	31	2	T06854	photosystem II pro
135	22	22.7	31	2	B82401	hypothetical prote	208	21	21.6	31	2	S32444	photosystem II pro
136	22	22.7	32	2	S23476	hypothetical prote	209	21	21.6	31	2	F82860	hypothetical prote
137	22	22.7	32	2	T14569	hypothetical prote	210	21	21.6	31	2	T07276	photosystem II pro
138	22	22.7	33	1	GN6PB	big gastrin [valid	211	21	21.6	32	2	S39628	probable urease (E
139	22	22.7	33	2	B29541	myosin - human [fr	212	21	21.6	32	2	S45670	Nef protein - huma
140	22	22.7	33	2	I61695	myosin - pig (frag	213	21	21.6	32	2	C70216	hypothetical prote
141	22	22.7	33	2	I46596	ycf12 protein - Ch	214	21	21.6	32	2	B85588	hypothetical prote
142	22	22.7	33	2	T08018	zona pellucida pro	215	21	21.6	32	2	S08482	regulatory protein
143	22	22.7	36	2	S35572	hypothetical prote	216	21	21.6	33	2	A70546	hypothetical prote
144	22	22.7	36	2	D82457	hypothetical prote	217	21	21.6	33	2	D70224	hypothetical prote
145	22	22.7	36	2	S34355	hypothetical prote	218	21	21.6	33	2	E82303	hypothetical prote
146	22	22.7	38	1	HSPY5	histone H5 - pigeo	219	21	21.6	33	2	E82391	hypothetical prote
147	22	22.7	38	2	S70800	Ivi protein III -	220	21	21.6	33	2	S88578	hypothetical prote
148	22	22.7	38	2	AB1966	photosystem II pro	221	21	21.6	34	2	A44038	leader peptide Tna
149	22	22.7	38	2	A10988	hypothetical prote	222	21	21.6	35	2	C95037	hypothetical prote
150	22	22.7	38	2	A44862	microtubule associ	223	21	21.6	35	2	A05057	hypothetical prote
151	22	22.7	39	2	S10315	photosystem II pro	224	21	21.6	35	2	T07509	photosystem II pro
152	22	22.7	39	2	S67938	hypothetical prote	225	21	21.6	35	2	T00254	hypothetical 4K pr
153	22	22.7	39	2	G82287	hypothetical prote	226	21	21.6	36	1	PRFG	pancreatic hormone
154	22	22.7	40	2	D61340	hypothetical prote	227	21	21.6	36	2	A28091	pancreatic hormone
155	22	22.7	40	2	AH2446	hypothetical prote	228	21	21.6	36	2	E69729	required for trans
156	22	22.7	40	2	S71301	ICL5 protein - Par	229	21	21.6	36	2	A69287	hypothetical prote
157	22	22.7	40	2	S44136	NADH2 dehydrogenas	230	21	21.6	37	2	C32021	bactericidin B-4 -
158	22	22.7	40	2	H82330	hypothetical prote	231	21	21.6	37	2	T07351	ribosomal protein
159	22	22.7	40	2	R64009	hypothetical prote	232	21	21.6	37	2	T07292	hypothetical prote
160	22	22.7	41	2	PQ0564	nonstructural prot	233	21	21.6	37	2	AH0844	photosystem II pro
161	22	22.7	41	2	PQ0565	nonstructural prot	234	21	21.6	38	1	F3KTL	photosystem II pro
162	22	22.7	41	2	T07263	hypothetical prote	235	21	21.6	38	2	S73157	photosystem II pro
163	22	22.7	42	2	A30010	cytochrome-c oxida	236	21	21.6	38	2	S88260	hypothetical prote
164	22	22.7	42	2	S01125	photosystem II pho	237	21	21.6	38	2	C70222	hypothetical prote
165	22	22.7	44	2	PQ0182	naringenin-chalcon	238	21	21.6	38	2	UN0418	hypothetical prote
166	22	22.7	44	2	T26893	hypothetical prote	239	21	21.6	38	2	S14141	hypothetical prote
167	22	22.7	45	2	H70249	hypothetical prote	240	21	21.6	39	2	I56173	BBV/C3d receptor -
168	22	22.7	45	2	H83936	hypothetical prote	241	21	21.6	39	2	C90523	hypothetical prote
169	22	22.7	46	2	D82163	hypothetical prote	242	21	21.6	39	2	C72426	hypothetical prote
170	22	22.7	46	2	G83936	hypothetical prote	243	21	21.6	39	2	S06901	hypothetical prote
171	22	22.7	47	2	B84357	50S ribosomal prot	244	21	21.6	39	2	D83811	hypothetical prote
172	22	22.7	47	2	A82385	hypothetical prote	245	21	21.6	40	2	A61320	plastocyanin - Ane
173	22	22.7	47	2	H83869	hypothetical prote	246	21	21.6	40	2	C61320	plastocyanin - Cle
174	22	22.7	48	1	PNCKGP	H+-transporting tw	247	21	21.6	40	2	E61320	plastocyanin - Era
175	22	22.7	48	2	B86779	hypothetical prote	248	21	21.6	40	2	PC4021	chaperonin contain

249	21	21.6	40	2	A40128	probable antigen 1	322	20	20.6	22	2	S55308	glutathione trans
250	21	21.6	40	2	D82440	hypothetical prote	323	20	20.6	23	2	S23518	stromelysin (EC 3.
251	21	21.6	40	2	S63506	protein B betaine-	324	20	20.6	23	2	I45916	epidermal keratin
252	21	21.6	40	2	F45095	photosystem I ligh	325	20	20.6	23	2	T50545	reductase (improte
253	21	21.6	40	2	T07472	hypothetical prote	326	20	20.6	23	2	B30988	hypothetical prote
254	21	21.6	41	2	D87311	hypothetical prote	327	20	20.6	23	2	B39313	chaperonin, 10K -
255	21	21.6	41	2	T07461	hypothetical prote	328	20	20.6	23	2	A83397	pyrroloquinoline q
256	21	21.6	41	2	B81565	hypothetical prote	329	20	20.6	24	2	D91209	tryptophanase lead
257	21	21.6	41	2	F82536	hypothetical prote	330	20	20.6	24	2	I54862	tnaC protein - Esc
258	21	21.6	41	2	AC3430	hypothetical prote	331	20	20.6	24	2	G86055	tryptophanase lead
259	21	21.6	42	2	T07285	hypothetical prote	332	20	20.6	24	2	G86055	T-cell receptor J-
260	21	21.6	42	2	S77750	probable ABC-type	333	20	20.6	25	2	S47204	1,2-alpha-D-manno
261	21	21.6	42	2	B85521	frame shifted GL2-	334	20	20.6	25	2	S30361	glutathione trans
262	21	21.6	42	4	S49071	cytochrome b559 co	335	20	20.6	25	2	S30361	glutathione trans
263	21	21.6	43	2	S78338	superoxide dismuta	336	20	20.6	26	2	A61499	T-cell receptor al
264	21	21.6	43	2	S53456	aidolase - Gnyeto	337	20	20.6	26	2	S58387	Ru-like protein HB
265	21	21.6	43	2	I77472	defensin - Pyrrhoc	338	20	20.6	26	2	A43768	homeotic protein H
266	21	21.6	43	2	S44463	Ig kappa chain V r	339	20	20.6	26	2	A43768	T-cell receptor va
267	21	21.6	43	2	S21085	Ig kappa chain V r	340	20	20.6	26	2	H44621	T-cell receptor H
268	21	21.6	43	2	PH0883	colicin 10 lysis p	341	20	20.6	27	2	H44621	pepsin A (EC 3.4.2
269	21	21.6	43	2	I41026	hypothetical prote	342	20	20.6	27	2	D49829	siatic acid-specif
270	21	21.6	43	2	G82207	hypothetical prote	343	20	20.6	28	2	PL0005	toxin-like protein
271	21	21.6	43	2	D97484	hypothetical prote	344	20	20.6	28	2	A46690	nicotinic acetylch
272	21	21.6	44	2	C37286	olfactory receptor	345	20	20.6	28	2	S06668	T-cell receptor be
273	21	21.6	44	2	F81079	hypothetical prote	346	20	20.6	28	2	S68643	photosystem II pro
274	21	21.6	44	2	I51889	prolactin receptor	347	20	20.6	29	2	S58390	photosystem II pro
275	21	21.6	44	2	T35464	hypothetical prote	348	20	20.6	29	2	S05032	hypothetical prote
276	21	21.6	45	2	C82233	hypothetical prote	349	20	20.6	29	2	F85570	hypothetical prote
277	21	21.6	45	2	G82155	hypothetical prote	350	20	20.6	30	2	A90027	G-alpha-1 protein
278	21	21.6	45	2	H82526	hypothetical prote	351	20	20.6	31	2	D70118	hypothetical prote
279	21	21.6	46	2	A30609	Ig kappa chain V-I	352	20	20.6	31	2	D70236	hypothetical prote
280	21	21.6	46	2	F69339	hypothetical prote	353	20	20.6	31	2	S76281	hypothetical prote
281	21	21.6	46	2	I47035	ubiquitin homolog	354	20	20.6	31	2	B60363	hypothetical prote
282	21	21.6	46	2	AG3569	hypothetical prote	355	20	20.6	31	2	G81558	apolipoprotein III
283	21	21.6	46	2	C81372	very hypothetical	356	20	20.6	32	2	C82768	hypothetical prote
284	21	21.6	47	2	PC4178	DNA-directed DNA p	357	20	20.6	32	4	A49016	flagellin, 24K - M
285	21	21.6	47	2	T05704	phosphoenolpyruvat	358	20	20.6	33	2	A61310	hypothetical prote
286	21	21.6	47	2	B69443	conserved hypothet	359	20	20.6	33	2	S50035	gag-jund mutant fu
287	21	21.6	47	2	F85877	hypothetical prote	360	20	20.6	34	2	S13439	nonhistone chromos
288	21	21.6	48	2	PH0884	Ig kappa chain V r	361	20	20.6	34	2	F81044	cytokine boGKO - b
289	21	21.6	48	2	S26117	outer membrane pro	362	20	20.6	34	2	F82458	lectin II, anti-H(
290	21	21.6	48	2	T13337	hypothetical prote	363	20	20.6	35	2	S07131	hypothetical prote
291	21	21.6	49	2	E95112	ribosomal protein	364	20	20.6	35	2	S43329	fatty-acid synthas
292	21	21.6	49	2	D97981	50S ribosomal prot	365	20	20.6	35	2	S68338	trypsin (EC 3.4.21
293	21	21.6	49	2	T07200	hypothetical prote	366	20	20.6	35	2	E48401	elastase inhibitor
294	21	21.6	49	2	E69428	LSU ribosomal prot	367	20	20.6	35	2	D82125	ribosomal protein
295	21	21.6	49	2	T02026	glycine-rich prote	368	20	20.6	35	2	S49309	hypothetical prote
296	21	21.6	49	2	T33471	hypothetical prote	369	20	20.6	35	2	A11823	oncofetal protein
297	21	21.6	49	2	S25829	serpin III - horse	370	20	20.6	36	1	CKAODP	photosystem II Pab
298	21	21.6	49	2	F5218	hypothetical prote	371	20	20.6	36	2	A05019	cecropin D - Chine
299	21	21.6	49	2	T12879	hypothetical prote	372	20	20.6	36	2	C54127	photosystem II pro
300	21	21.6	49	2	C97061	hypothetical prote	373	20	20.6	36	2	S43284	dolichyl-diphospho
301	21	21.6	50	1	INFIS	insulin - shorthor	374	20	20.6	36	2	A95135	gallinacin - chick
302	21	21.6	50	2	S67520	ferredoxin (3Fe-4S	375	20	20.6	37	2	E53578	hypothetical prote
303	21	21.6	50	2	C69094	ribosomal protein	376	20	20.6	37	2	A32021	esculentin 2a - ed
304	21	21.6	50	2	G97836	hypothetical prote	377	20	20.6	37	2	B32021	esculentin 2b - ed
305	21	21.6	50	2	G90974	hypothetical prote	378	20	20.6	37	2	A82057	bactericidin B-2 -
306	21	21.6	50	2	H85821	unknown protein en	379	20	20.6	37	2	D83199	bactericidin B-3 -
307	20.5	21.1	33	2	A46116	penton base protei	380	20	20.6	38	1	R5EC36	hypothetical prote
308	20.5	21.1	33	2	A83591	hypothetical prote	381	20	20.6	38	1	R5EC36	50S ribosomal chai
309	20.5	21.1	39	2	F82329	hypothetical prote	382	20	20.6	38	2	H83113	50S ribosomal prot
310	20.5	21.1	43	1	TNBOB4	thymosin beta-4 -	383	20	20.6	38	2	D91149	50S ribosomal subu
311	20.5	21.1	43	1	A38054	tubulin beta chain	384	20	20.6	38	2	H85994	50S ribosomal subu
312	20.5	21.1	44	1	A38682	thymosin beta-4 [v	385	20	20.6	38	2	AG0028	50S ribosomal prot
313	20.5	21.1	44	1	JQ1489	thymosin beta-4 -	386	20	20.6	38	2	T11950	Photosystem II pro
314	20.5	21.1	44	1	TNRBT4	thymosin beta-4 -	387	20	20.6	38	2	S78330	photosystem II pro
315	20.5	21.1	45	2	JCS274	thymosin beta - hu	388	20	20.6	38	2	S76107	photosystem II ligh
316	20.5	21.1	50	1	A37217	thymosin beta-4 -	389	20	20.6	38	2	G45095	kallikrein - mouse
317	20	20.6	20	2	S00494	hemocyanin chain I	390	20	20.6	38	2	I70018	hypothetical prote
318	20	20.6	20	2	S00493	hemocyanin chain I	391	20	20.6	38	2	D42227	hypothetical prote
319	20	20.6	20	2	S33867	quinaldic acid 4-o	392	20	20.6	38	2	G82102	hypothetical prote
320	20	20.6	20	2	A47687	3-dehydroquinat d	393	20	20.6	38	2	E82424	hypothetical prote
321	20	20.6	21	2	S47189	T-cell receptor J-	394	20	20.6	38	2	E82424	hypothetical prote

R;White, K.C.; Babbitt, P.C.; Buechter, D.D.; Kenyon, G.L.

U. Protein Chem. 11, 489-494, 1992

A;Title: The principal islet of the Coho salmon (*Oncorhynchus kisutch*) contains the BB is

A;Reference number: A53875; MUID:93080727; PMID:1449598

A;Accession: A53875

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-20 <WHI>

A;Cross-references: UNIPROT:Q9PSI5

A;Experimental source: Brockmann body, principal islet

A;Note: sequence extracted from NCBI backbone (NCBIP:120599)

C;Superfamily: creatine kinase; creatine kinase repeat homology

C;Keywords: phosphotransferase

Query Match 34.0%; Score 33; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 35;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HLNSKIAFKIVSQEP 18

Db 6 HNNFKLNFKVVEEP 20

RESULT 3

D82660

hypothetical protein XF1612 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: D82660

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: D82660

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-50 <SIM>

A;Cross-references: UNIPROT:Q9PCZ2; GB:AE003988; GB:AE003849; NID:g9106653; PIDN:AAF8442

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.Y.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF1612

Query Match 34.0%; Score 33; DB 2; Length 50;

Best Local Similarity 46.7%; Pred. No. 98;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQEPA 19

Db 23 LNKLSLNKIKNEEA 37

RESULT 4

B57473

caveolin beta form - mouse (fragments)

C;Species: Mus musculus (house mouse)

C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 11-Jan-2000

C;Accession: B57473

R;Scherer, P.E.; Tang, Z.; Chun, M.; Sargiacomo, M.; Lodish, H.F.; Lisanti, M.P.

J. Biol. Chem. 270, 16395-16401, 1995

A;Title: Caveolin isoforms differ in their N-terminal protein sequence and subcellular

A;Reference number: A57473; MUID:95332353; PMID:7608210

A;Accession: B57473

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-44 <SCH>

C;Superfamily: caveolin

C;Keywords: alternative initiators; blocked amino end

Query Match 31.4%; Score 30.5; DB 2; Length 44;

Best Local Similarity 50.0%; Pred. No. 2.3e+02;

Matches 9; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

Qy 4 HLNS---KIAPKIVSQEP 18

Db 13 HLNDVVVKIDFEDVIAEP 30

RESULT 5

JC6158

viral capsid protein 47 - human herpesvirus 8

C;Species: human herpesvirus 8, Kaposi sarcoma-associated virus

C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004

C;Accession: JC6158

R;Zhong, W.; Wang, H.; Herndier, B.; Ganew, D.

Proc. Natl. Acad. Sci. U.S.A. 93, 6641-6646, 1996

A;Title: Restricted expression of Kaposi sarcoma-associated herpesvirus (human herpesvi

A;Reference number: JC6153; MUID:96270595; PMID:8692871

A;Accession: JC6158

A;Molecule type: mRNA

A;Residues: 1-47 <ZHO>

A;Cross-references: UNIPROT:Q98156; GB:U75698

C;Comment: This protein is a marker for the subpopulation involved in productive infect

C;Keywords: membrane protein

Query Match 30.9%; Score 30; DB 2; Length 47;

Best Local Similarity 55.6%; Pred. No. 3.1e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PNHLNSKIA 10

Db 27 PNHTNGTMA 35

RESULT 6

H82096

hypothetical protein VC2284 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: H82096

R;Heddelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers,

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: H82096

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-33 <HEI>

A;Cross-references: UNIPROT:Q9KPS8; GB:AE004299; GB:AE003852; NID:g9656835; PIDN:AAF954

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC2284

A;Map position: 1

Query Match 29.9%; Score 29; DB 2; Length 33;

Best Local Similarity 30.8%; Pred. No. 3.1e+02;

Matches 4; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EPNHLNSKIAPKI 13

Db 1 EPNHLNSKIAPKI 13

```

Db      21 DSNQISNDVSPFKI 33

RESULT 7
Hypothetical protein VCA0226 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82484
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82484
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-40 <HEI>
A:Cross-references: UNIPROT:Q9KMU3; GB:AE004363; GB:AE003853; NID:g9657618; PIDN:AAF9613
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0226
A:Map position: 2

Query Match      29.9%; Score 29; DB 2; Length 40;
Best Local Similarity 45.5%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 PNLNSKIAPK 12
      ||||| :
Db      30 PNHLELVGGK 40

RESULT 8
Hypothetical protein CC1197 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87397
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87397
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-49 <STO>
A:Cross-references: UNIPROT:Q9A828; GB:AE005673; NID:gl3422522; PIDN:AAK23180.1; GSPDB:C
C:Genetics:
A:Gene: CC1197

Query Match      29.4%; Score 28.5; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY      3 NNLNSKI-AFKIVS 15
      ||| :| :|
Db      18 NHLRTEVTRFLIVS 31

RESULT 9
Hypothetical protein 1 - yeast (Williopsis suaveolens)
C:Species: Williopsis suaveolens
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S47281
R:Nosek, J.
submitted to the EMBL Data Library, February 1994
A:Reference number: S47281
A:Accession: S47281

Db      21 DSNQISNDVSPFKI 33

```

```

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-24 <NOS>
A:Cross-references: UNIPROT:Q01228; EMBL:X77691; NID:g535203; PID:g535204
C:Superfamily: yeast mitochondrion oxii 3' region hypothetical protein

Query Match      28.9%; Score 28; DB 2; Length 24;
Best Local Similarity 38.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      3 NNLNSKIAPKIVS 15
      :: :| :| :|
Db      12 SNTNNKILFRLES 24

RESULT 10
Hypothetical protein SP1629 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: D95189
R:Fettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: D95189
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-34 <KUR>
A:Cross-references: UNIPROT:Q97PI6; GB:AE005672; PIDN:AAK75709.1; PID:gl4973118; GSPDB:
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1629

Query Match      28.9%; Score 28; DB 2; Length 34;
Best Local Similarity 40.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      4 HLNSKIAPKIVSQEP 18
      : :| :| :|
Db      20 HILQXMKNKLSQNP 34

RESULT 11
SI7507
cytokine - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: SI7507
R:Joese, P.J.; Collins, P.D.; Perkins, J.A.; Beaubien, B.C.; Totty, N.F.; Waterfield, M.
Biochem. J. 278, 493-497, 1991
A:Title: Identification of a second neutrophil-chemoattractant cytokine generated durin
d structural relationship to melanoma-growth-stimulatory activity.
A:Reference number: SI7507; MUID:91378900; PMID:1898341
A:Accession: SI7507
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-36 <JOS>
A:Cross-references: UNIPROT:Q28724
C:Superfamily: beta-thromboglobulin

Query Match      28.9%; Score 28; DB 2; Length 36;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      4 HLNSKIAPKIVSQEP 18
      ||| :| :| :|
Db      17 HLKSIQSLKVLSPGP 31

```

```

RESULT 12
PQ0444
hypothetical protein 30 - slime mold (Dictyostelium giganteum) plasmid Dgpl (fragment)
C:Species: Dictyostelium giganteum
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PQ0444
R:Yin, Y.; Welker, D.L.
Plasmid 28, 37-45, 1992
A:Title: Dictyostelium giganteum plasmid Dgpl is a member of the Ddp2 plasmid family.
A:Reference number: PQ0444; NCID:92390516; PMID:1518911
A:Accession: PQ0444
A:Molecule type: DNA
A:Residues: 1-30 <YIN>
A:Cross-references: UNIPROT:Q26257; GB:S43953; NID:G254926; PIDN:AAB23143.1; PID:G254927
A:Experimental source: strain Dg61
C:Genetics:
A:Genome: plasmid

Query Match      27.8%; Score 27; DB 2; Length 30;
Best Local Similarity 27.3%; Pred. No. 6.2e+02;
Matches 3; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHNSKIAF 11
   :|:::|:|
DB 2 DPNIDHKVGF 12

RESULT 13
S22603
ribosomal protein L26, cytosolic - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S22603
R:Lorenzi, F.; Francesconi, A.; Jappelli, R.; Amaldi, F.
Nucleic Acids Res. 20, 1859-1863, 1992
A:Title: Analysis of mRNAs under translational control during Xenopus embryogenesis: iso
A:Reference number: S22601; MUID:92253404; PMID:1579486
A:Accession: S22603
A:Molecule type: mRNA
A:Residues: 1-33 <LOR>
A:Cross-references: UNIPROT:P49629; EMBL:X64211; NID:G65086; PIDN:CAB56811.1; PID:G60068
C:Superfamily: rat ribosomal protein L26
C:Keywords: cytosol; protein biosynthesis; ribosome

Query Match      27.8%; Score 27; DB 2; Length 33;
Best Local Similarity 37.5%; Pred. No. 6.9e+02;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 PNLNSKIAFKIVSQE 17
   :|:|:|:|
DB 1 PSHVRKIMSWPLSKE 16

RESULT 14
A61235
fibroblast-activating factor 32K precursor - human (fragments)
C:Species: Homo sapiens (man)
C>Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 09-Jul-2004
C:Accession: A61235
R:Demeter, J.; Medzihradsky, D.; Kha, H.; Goetzl, E.J.; Turck, C.W.
Immunology 72, 350-354, 1991
A:Title: Isolation and partial characterization of the structures of fibroblast activati
A:Reference number: A61235; MUID:91224664; PMID:2026444
A:Accession: A61235
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-36 <DEM>
A:Cross-references: UNIPROT:Q7M4R4

Query Match      27.8%; Score 27; DB 2; Length 36;
Best Local Similarity 33.3%; Pred. No. 7.5e+02;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 6 NSKIAFKIV 14
   |:::|:|
DB 24 NDRVAIKVL 32

RESULT 15
G70164
hypothetical protein BB0520 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: G70164
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: G70164
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-38 <KLE>
A:Cross-references: UNIPROT:O51470; GB:AE001154; GB:AE000783; NID:G2688431; PIDN:AAC668
A:Experimental source: strain B31

Query Match      27.8%; Score 27; DB 2; Length 38;
Best Local Similarity 38.5%; Pred. No. 8e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQE 17
   :|:|:|:|
DB 20 INKIKFILTCK 32

RESULT 16
S70031
placental lactogen - sheep (fragments)
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 09-Jul-2004
C:Accession: S70031
R:Cymes, G.D.; Wolfenstein-Todel, C.
Biochim. Biophys. Acta 1294, 31-36, 1996
A:Title: Identification of a tyrosine residue in ovine placental lactogen as essential
A:Reference number: S70031; MUID:96219883; PMID:8639711
A:Accession: S70031
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-12;13-20;21-30;31-35;36-43 <CYM>
A:Cross-references: UNIPROT:P16038

Query Match      27.8%; Score 27; DB 2; Length 43;
Best Local Similarity 83.3%; Pred. No. 9.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 NSKIAF 11
   |||:|
DB 8 NSKLAF 13

RESULT 17
E72258
hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: E72258
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: E72258
A>Status: preliminary

```

A:Molecule type: DNA
A:Residues: 1-47 <ARN>
A:Cross-references: UNIPROT:Q9X1D2; GB:AE001793; GB:AE000512; NID:g4981963; PIDN:AAD3648
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1412

Query Match 27.8%; Score 27; DB 2; Length 47;
Best Local Similarity 38.5%; Pred. No. 1e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKI 13
| : | | | : | : |
Db 2 EPHLLNSVVCYKL 14
| : | | | : | : |

RESULT 18
S62189
phosphatidylserine decarboxylase homolog - Azotobacter vinelandii (fragment)
C:Species: Azotobacter vinelandii
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S62189
R:Colnaghi, R.; Pagani, S.; Kennedy, C.; Drummond, M.
Eur. J. Biochem. 236, 240-248, 1996
A:Title: Cloning, sequence analysis and overexpression of the rhodanese gene of Azotobacter vinelandii
A:Reference number: S62187; MUID:96184904; PMID:8617271
A:Accession: S62189
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-48 <COL>
A:Cross-references: UNIPROT:Q44558; EMBL:L42346; NID:g1069990; PIDN:AAB03238.1; PID:g139

Query Match 27.8%; Score 27; DB 2; Length 48;
Best Local Similarity 41.7%; Pred. No. 1e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNHLSKIAFKI 13
| : | | | : | : |
Db 14 PHHLSRAAGRL 25
| : | | | : | : |

RESULT 19
S70246
nifW2 protein - Anabaena variabilis (fragment)
C:Species: Anabaena variabilis
C>Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C:Accession: S70246; S49993
R:Schrautemeier, B.; Neveling, U.; Schmitz, S.
Mol. Microbiol. 18, 357-369, 1995
A:Title: Distinct and differently regulated Mo-dependent nitrogen-fixing systems evolved in Anabaena variabilis
A:Reference number: S70242; MUID:96296457; PMID:8709854
A:Accession: S70246
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-34 <SCH>
A:Cross-references: UNIPROT:P46054; EMBL:Z46890; NID:g599647; PIDN:CRA86988.1; PID:g5996
A:Experimental source: ATCC 29413
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994
C:Genetics:
A:Gene: nifW2

Query Match 27.3%; Score 26.5; DB 2; Length 34;
Best Local Similarity 42.9%; Pred. No. 8.7e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 5 LNSKIAFKIVSQEP 18
| : | : | : | : | : |
Db 10 LEQKL-FKVFQKQP 22
| : | : | : | : | : |

hypothetical protein XF1790 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82636
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: D82636
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-37 <SIM>
A:Cross-references: UNIPROT:O9PCI9; GB:AE004001; GB:AE003849; NID:g9106864; PIDN:AAF845
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, E.L.P.; Camargo, A.A.; Camargo, D.M.; Carraro, D.M.; Carrier, as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froberg, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig, Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawada, A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva, M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1790

Query Match 27.3%; Score 26.5; DB 2; Length 37;
Best Local Similarity 41.2%; Pred. No. 9.5e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 2 PNHLSKIAFKIVSQEP 18
| : | : | : | : | : |
Db 15 PNLVNGQIAC-VLSGSP 30
| : | : | : | : | : |

RESULT 21
S78355
hypothetical protein 27 - Odontella sinensis chloroplast
C:Species: chloroplast Odontella sinensis
C>Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 09-Jul-2004
C:Accession: S78355
R:Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
Plant Mol. Biol. Rep. 13, 336-342, 1995
A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sinensis
A:Reference number: S78238
A:Accession: S78355
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-27 <KOW>
A:Cross-references: UNIPROT:P49836; EMBL:Z67753; NID:g1185127; PID:e212176; PID:g118524
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
C:Genetics:
A:Gene: chloroplast
C:Keywords: chloroplast

Query Match 26.8%; Score 26; DB 2; Length 27;
Best Local Similarity 46.7%; Pred. No. 8.2e+02;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
| : | : | : | : | : |
Db 8 LRKKIAEVWIFQNP 22
| : | : | : | : | : |

RESULT 22
E82446
hypothetical protein VCA0541 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: E82446
 R;Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A/Reference number: A82035; MUID:20406833; PMID:10952301
 A/Accession: E82446
 A>Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-31 <HEI>
 A/Cross-references: UNIPROT:Q9KM48; GB:AE004385; GB:AE003853; NID:G9657946; PIDN:AAF9644
 A/Experimental source: serogroup O1; strain N16961; biotype El Tor
 C/Genetics:
 A/Gene: VCA0541
 A/Map position: 2

Query Match 26.8%; Score 26; DB 2; Length 31;
 Best Local Similarity 80.0%; Pred. No. 9.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NHLNS 7
 |||
 Db 6 NHINS 10

RESULT 23

JQ1620

hypothetical 4.3K protein - human immunodeficiency virus type 1
 C/Species: human immunodeficiency virus type 1, HIV-1
 C/Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
 C/Accession: JQ1620
 R;Smith, J.; Azad, A.; Deacon, N.
 J. Gen. Virol. 73, 1825-1828, 1992
 A>Title: Identification of two novel human immunodeficiency virus type 1 splice acceptor sites in HIV-1
 A/Reference number: JQ1620; MUID:92333269; PMID:1629704
 A/Accession: JQ1620
 A>Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-34 <SML>
 A/Cross-references: UNIPROT:Q9Q5D3; UNIPROT:Q9Q5D2; UNIPROT:Q71959; UNIPROT:Q71985; UNIPROT:Q9FXW9; UNIPROT:Q71964; UNIPROT:Q9Q594; UNIPROT:Q71960; UNIPROT:Q40177; UNIPROT:Q9WLM4; UNIPROT:Q9PXW8; UNIPROT:Q9Q595; UNIPROT:Q9PY48; UNIPROT:Q9WM24; UNIPROT:Q9Q5G8
 C/Superfamily: AIDS nef protein

Query Match 26.8%; Score 26; DB 2; Length 34;
 Best Local Similarity 41.7%; Pred. No. 1.1e+03;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NSKIAFKIVSQE 17
 :||:| | | | |
 Db 14 DSRLEAFHHVARE 25

RESULT 24

F82051

hypothetical protein VC2639 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
 C/Species: *Vibrio cholerae*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: F82051
 R;Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A/Reference number: A82035; MUID:20406833; PMID:10952301
 A/Accession: F82051
 A>Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-35 <HEI>

A/Cross-references: UNIPROT:Q9KNU1; GB:AE004330; GB:AE003852; NID:G9657225; PIDN:AAF9578

A/Experimental source: serogroup O1; strain N16961; biotype El Tor
 C/Genetics:
 A/Gene: VC2639
 A/Map position: 1

Query Match 26.8%; Score 26; DB 2; Length 35;
 Best Local Similarity 42.9%; Pred. No. 1.1e+03;
 Matches 6; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

Qy 3 NNLNSKIAFKIVSQ 16
 ||||| | | | |
 Db 2 NELNSK--FNLIQK 13

RESULT 25

S21594

Ig heavy chain V region (clone 16B8) - mouse (fragment)
 C/Species: *Mus musculus* (house mouse)
 C/Date: 20-Feb-1995 #sequence_revision 26-Apr-1996 #text_change 30-May-1997
 C/Accession: S21594
 R;Kaartinen, M.

submitted to the EMBL Data Library, May 1992

A/Reference number: S21591
 A/Accession: S21594
 A/Molecule type: mRNA
 A/Residues: 1-36 <KAA>
 A/Cross-references: EMBL:X66461
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotrimer; immunoglobulin

Query Match 26.8%; Score 26; DB 2; Length 36;
 Best Local Similarity 41.7%; Pred. No. 1.1e+03;
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 7 SKIAFKIVSQEP 18
 :||:| | | | |
 Db 2 SQVFFKWNLSLOP 13

RESULT 26

A60867

properdin - rabbit (fragment)
 N/Alternate names: factor P
 C/Species: *Oryctolagus cuniculus* (domestic rabbit)
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C/Accession: A60867
 R;Nakano, Y.; Matsuda, T.; Sakamoto, T.; Tomita, M.
 J. Immunol. Methods 90, 77-83, 1986
 A>Title: Isolation and characterization of rabbit properdin of the alternative complement

A/Reference number: A60867; MUID:86225592; PMID:3635564

A/Accession: A60867
 A/Molecule type: protein
 A/Residues: 1-38 <NAK>
 A/Cross-references: UNIPROT:Q7M2T4
 C/Superfamily: human properdin precursor; thrombospondin type 1 repeat homology
 C/Keywords: complement alternate pathway; plasma

Query Match 26.8%; Score 26; DB 2; Length 38;
 Best Local Similarity 50.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EPNHNSKIAFK 12
 ||||| | | | |
 Db 27 EPCCLNAAVAFQ 38

RESULT 27

H70246

hypothetical protein BSJ15 - Lyme disease spirochete plasmid J/lp38
 C/Species: *Borrelia burgdorferi* (Lyme disease spirochete)
 C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 C/Accession: H70246
 R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

```

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: H70246
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-42 <KLE>
A:Cross-references: UNIPROT:O50770; GB:AE000787; NID:92690175; PIDN:AAC66101.1; PID:9269
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match      26.8%; Score 26; DB 2; Length 42;
Best Local Similarity 62.5%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 9 IAFKIVSQ 16
    |||:|:|
Db 8 IIFKFIQS 15

RESULT 28
I36946
Haptoglobin pseudogene Hpr - chimpanzee (fragment)
C:Species: Pan troglodytes (Chimpanzee)
C:Date: 15-Feb-1996 #sequence_revision 15-Feb-1996 #text_change 31-May-1996
C:Accession: I36946
R:McEvoy, S.; Maeda, N.
J. Biol. Chem. 263, 15740-15747, 1988
A:Title: Complex events in the evolution of the haptoglobin gene cluster in primates.
A:Reference number: I36945; MUID:89008487; PMID:3170608
A:Accession: I36946
A:Status: conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-42 <RES>
A:Cross-references: GB:M20761; GB:J04045
C:Genetics:
A:Gene: Hpr
C:Keywords: pseudogene

Query Match      26.8%; Score 26; DB 4; Length 42;
Best Local Similarity 38.9%; Pred. No. 1.3e+03;
Matches 7; Conservative 5; Mismatches 2; Indels 4; Gaps 1;

QY 4 HLNSKIAP----KIVSQE 17
    |||:|:|:|:|:|
Db 18 HLDAGKSPFWQAKMVS HQ 35

RESULT 29
E82428
Hypothetical protein VCA0701 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82428
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: E82428
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-43 <HEI>
A:Cross-references: UNIPROT:Q9KLP2; GB:AE004399; GB:AE003853; NID:99658111; PIDN:AAF9660
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0701
A:Map position: 2

```

```

Query Match      26.8%; Score 26; DB 2; Length 43;
Best Local Similarity 35.7%; Pred. No. 1.4e+03;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEP 18
    |||:|:|
Db 24 LYSKVLRKILSEP 37

RESULT 30
A19434
Hypothetical protein 1 - Elaphe radiata (fragment)
C:Species: Elaphe radiata
C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 09-Jul-2004
C:Accession: A19434
R:Epplen, J.T.; McCarrey, J.R.; Sutou, S.; Ohno, S.
Proc. Natl. Acad. Sci. U.S.A. 79, 3798-3802, 1982
A:Title: Base sequence of a cloned snake W-chromosome DNA fragment and identification o
A:Reference number: A19434; MUID:82247938; PMID:6954524
A:Accession: A19434
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-44 <EPP>
A:Cross-references: UNIPROT:Q7LZ30

Query Match      26.8%; Score 26; DB 2; Length 44;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLNSK 8
    |||:|
Db 11 PSHLKS R 17

RESULT 31
S66648
major acute phase protein - Chilean potato-tree (fragments)
C:Species: Solanum crispum (Chilean potato-tree)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S66648
R:Gonzalez-Ramon, N.; Alava, M.A.; Sarsa, J.A.; Pineiro, M.; Escartin, A.; Garcia-Gil,
FBBS Lett. 371, 227-230, 1995
A:Title: The major acute phase serum protein in pigs is homologous to human plasma kall
A:Reference number: S66648; MUID:96013138; PMID:7556597
A:Accession: S66648
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-27;28-45 <GON>
A:Cross-references: UNIPROT:Q7M214
C:Superfamily: inter-alpha-trypsin inhibitor complex component II

Query Match      26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 28.6%; Pred. No. 1.4e+03;
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEP 18
    |||:|:|
Db 16 VSRFPASKVVTSSP 29

RESULT 32
AC1020
Hypothetical protein STY4474 [imported] - Salmonella enterica subsp. enterica serovar Typhi
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC1020
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrah
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K

```


A>Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AC1020
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-48 <PAR>
 A:Cross-references: GB:AU513382; PIDN:CAD09260.1; PID:gi6505264; GSPDB:GN00176
 A:Genetics:
 A:Gene: STY4474

Query Match 26.8%; Score 26; DB 2; Length 48;
 Best Local Similarity 62.5%; Pred. No. 1.6e+03;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 NSKIAPKI 13
 |||:|
 DB 24 NSKVIFYI 31

RESULT 33
 138223
 protein-serine/threonine kinase - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004
 C:Accession: 138223; S37425
 R:Schultz, S.J.; Nigg, E.A.
 Cell Growth Differ. 4, 821-830, 1993
 A>Title: Identification of 21 novel human protein kinases, including 3 members of a family
 A:Reference number: 138211; MUID:94100173; PMID:8274451
 A:Accession: 138223
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-48 <RES>
 A:Cross-references: UNIPROT:Q15455; EMBL:Z25433; NID:g405744; PIDN:CAA80920.1; PID:g40574

Query Match 26.8%; Score 26; DB 2; Length 48;
 Best Local Similarity 44.4%; Pred. No. 1.6e+03;
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLNSKIA 10
 |||:|
 DB 25 PNYISPEIA 33

RESULT 34
 E81927
 hypothetical protein NMA0824 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup
 C:Species: *Neisseria meningitidis*
 C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
 C:Accession: E81927
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A>Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: E81927
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-50 <PAR>
 A:Cross-references: UNIPROT:Q9UVI7; GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB8410
 A:Experimental source: serogroup A, strain Z2491
 A:Genetics:
 A:Gene: NMA0822; NMA0824

Query Match 26.8%; Score 26; DB 2; Length 50;
 Best Local Similarity 66.7%; Pred. No. 1.6e+03;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 10 AFKIVSQEP 18
 |||:|
 DB 15 AFKPVLRQP 23

RESULT 35

WMBPPI
 gene i protein - phage PRD1
 C:Species: phage PRD1
 A>Note: host *Escherichia coli*
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
 C:Accession: A36776
 R:Bamford, J.K.H.; Haenninen, A.L.; Pakula, T.M.; Ojala, P.M.; Kalkkinen, N.; Frilander
 Virolgy 183, 658-676, 1991
 A>Title: Genome organization of membrane-containing bacteriophage PRD1.
 A:Reference number: A40477; MUID:91306449; PMID:1853567
 A:Accession: A36776
 A:Molecule type: DNA
 A:Residues: 1-42 <BAM>
 A:Cross-references: UNIPROT:P27386; GB:M69077; NID:g215765; PIDN:AAA32464.1; PID:g21577
 A:Genetics:
 A:Gene: i
 C:Superfamily: phage PRD1 gene i protein

Query Match 26.3%; Score 25.5; DB 1; Length 42;
 Best Local Similarity 42.9%; Pred. No. 1.6e+03;
 Matches 6; Conservative 4; Mismatches 1; Indels 3; Gaps 1;

QY 9 IAPKIV--SOEPA 19
 :||:|
 DB 29 LAAKVINRAQEP 42

RESULT 36
 S08605
 hypothetical protein 1 estrogen receptor 5'-region - chicken
 C:Species: *Gallus gallus* (chicken)
 C>Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 16-Aug-2004
 C:Accession: S08605
 R:Krust, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.; Bornert, J.M.; Chambon, P.
 EMBO J. 5, 891-897, 1986
 A>Title: The chicken oestrogen receptor sequence: homology with v-erbA and the human oe
 A:Reference number: S07192; MUID:86247578; PMID:3755102
 A:Accession: S08605
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-20 <KRU>
 A:Cross-references: EMBL:X03805; NID:g63378; PIDN:CAA27431.1; PID:g63379
 A>Note: the authors translated the codon TTT for residue 5 as Gly and TTC for residue 14

Query Match 25.8%; Score 25; DB 2; Length 20;
 Best Local Similarity 50.0%; Pred. No. 8.7e+02;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
 |::||
 DB 5 FNMISAE 12

RESULT 37
 B46174
 RNA-binding protein TIAR - human (fragment)
 N:Alternate names: cytotoxic granule-associated RNA-binding protein; TIA-1 related prote
 C:Species: Homo sapiens (man)
 C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Jul-1998
 C:Accession: B46174
 R:Kawakami, A.; Tian, Q.; Duan, X.; Streuli, M.; Schlosseman, S.F.; Anderson, P.
 Proc. Natl. Acad. Sci. U.S.A. 89, 8681-8685, 1992
 A>Title: Identification and functional characterization of a TIA-1-related nucleolysin.
 A:Reference number: A46174; MUID:92409580; PMID:1326761
 A:Accession: B46174
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-20 <KAW>
 A>Note: sequence extracted from NCBI backbone (NCBI:114067, NCBIP:114068)
 C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprote

Query Match 25.8%; Score 25; DB 2; Length 20;

Best Local Similarity 21.4%; Pred. No. 8.7e+02;
Matches 3; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIV 14
: ||| : : :
Db 5 DSRVNSVGVPSVL 18

RESULT 38

JU0386
nitrile hydratase (EC 4.2.1.84) - Acinetobacter sp. (fragment)

N;Alternate names: nitrilase

C;Species: Acinetobacter sp.

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 07-May-1999

C;Accession: JU0386

R;Yamamoto, K.; Komatsu, K.

Agric. Biol. Chem. 55, 1459-1466, 1991

A;Title: Purification and characterization of nitrilase responsible for the enantioselect

A;Reference number: JU0386; MUID:91345837; PMID:1369128

A;Accession: JU0386

A;Molecule type: protein

A;Residues: 1-21 <YAM>

A;Experimental source: strain AK226

C;Comment: The enzyme prefers S-(-)-2-(4'-isobutylphenyl)-propionitrile to R-(+)-2-(4'-i

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 25.8%; Score 25; DB 2; Length 21;
Best Local Similarity 46.2%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQEP 18
: ||| : : :
Db 4 NSKFLAATVQAE 16

RESULT 39

B46451
RNA-directed RNA polymerase (EC 2.7.7.48) - parainfluenza virus type 3 (strain 47885) (f

C;Species: parainfluenza virus type 3

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: B46451

R;Storey, D.G.; Cote, M.J.; Dimock, K.; Kang, C.Y.

Intervirology 27, 69-80, 1987

A;Title: Nucleotide sequence of the coding and flanking regions of the human parainfluen

A;Reference number: A46451; MUID:88032139; PMID:2822598

A;Accession: B46451

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-26 <STO>

A;Cross-references: UNIPROT:P12577; GB:M20402; NID:G332720; PIDN:AAA46857.1; PID:G332722

C;Superfamily: parainfluenza virus RNA-directed RNA polymerase

C;Keywords: ATP; nucleotidyltransferase

Query Match 25.8%; Score 25; DB 2; Length 26;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 HLSNKI 9
: ||| : : :
Db 19 HLSNPI 24

RESULT 40

A22977

delta-endotoxin - Bacillus thuringiensis (fragment)

C;Species: Bacillus thuringiensis

C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 07-Feb-1997

C;Accession: A22977

R;Armstrong, J.L.; Rohmann, G.F.; Beaudreau, G.S.

J. Bacteriol. 161, 39-46, 1985

A;Reference number: A22977; MUID:85104736; PMID:2981808

A;Accession: A22977

A;Molecule type: protein

A;Residues: 1-30 <ARM>

C;Superfamily: 28K parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 25.8%; Score 25; DB 2; Length 30;
Best Local Similarity 30.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIA 10
: ||| : : :
Db 4 DPNEINLLS 13

RESULT 41

A69827
hypothetical protein yhdS - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C;Accession: A69827

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bette

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: A69827

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-36 <KUN>

A;Cross-references: UNIPROT:O07588; GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12797

A;Experimental source: strain 168

C;Genetics:

A;Gene: yhdS

Query Match 25.8%; Score 25; DB 2; Length 36;
Best Local Similarity 35.7%; Pred. No. 1.7e+03;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 PNHLSKIAFKIVS 15
: ||| : : :
Db 9 PHHTTSKQSLLIN 22

RESULT 42

D82617

hypothetical protein XP1962 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: D82617

R.anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: D82617

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-36 <STM>

A;Cross-references: UNIPROT:Q9PC23; GB:AE004015; GB:AE003849; NID:G9107057; PIDN:AAF84

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

A;Description: One hundred R2R3-MYB genes in the genome of Arabidopsis thaliana.
 A;Reference number: Z25968
 A;Accession: T52125
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-45 <ROM>
 A;Cross-references: UNIPROT:O65898; UNIPROT:Q9FR97; EMBL:Z95778; PIDN:CAB09210.1
 A;Accession: T52126
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-45 <RO2>
 A;Cross-references: EMBL:Z95788; PIDN:CAB09220.1
 C;Genetics:
 A;Gene: MYB56; MYB65

Query Match 25.8%; Score 25; DB 2; Length 45;
 Best Local Similarity 40.0%; Pred. No. 2.2e+03;
 Matches 6; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQE 17
 ||||| : |||||
 DB 5 NHLRNLKKGAFSQE 19

RESULT 48

S78243
 hypothetical protein 46 - Odontella sinensis chloroplast
 C;Species: chloroplast Odontella sinensis
 C;Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 09-Jul-2004
 C;Accession: S78243
 R;Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
 Plant Mol. Biol. Rep. 13, 336-342, 1995
 A;Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sinensis
 A;Reference number: S78238
 A;Accession: S78243
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-46 <KOW>
 A;Cross-references: UNIPROT:P49828; EMBL:Z67753; NID:g1185127; PID:e211809; PID:g1185133
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
 C;Genetics:
 A;Genome: chloroplast
 C;Keywords: chloroplast

Query Match 25.8%; Score 25; DB 2; Length 46;
 Best Local Similarity 36.8%; Pred. No. 2.2e+03;
 Matches 7; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 1 BPNHLNS--KIAFKIVSQE 17
 : || || || : || || :
 DB 22 KPYSLSNTTKTLKIKNK 40

RESULT 49

T05086
 hypothetical protein T6K21.180 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C;Accession: T05086
 R;Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Bancroft, I.; Mewes, submitted to the Protein Sequence Database, February 1998
 A;Reference number: Z15397
 A;Accession: T05086
 A;Molecule type: DNA
 A;Residues: 1-46 <BEV>
 A;Cross-references: UNIPROT:O49699; EMBL:AL021889
 A;Experimental source: cultivar Columbia; BAC clone T6K21
 C;Genetics:
 A;Map position: 4
 A;Introns: 38/3
 A;Note: T6K21.180

Query Match 25.8%; Score 25; DB 2; Length 46;

Best Local Similarity 66.7%; Pred. No. 2.2e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLNS 7
 || : ||
 DB 20 PNEINS 25

RESULT 50

S31005
 gene 60 protein - Mycobacterium phage L5
 C;Species: Mycobacterium phage L5
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C;Accession: S31005
 R;Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
 Mol. Microbiol. 7, 407-417, 1993
 A;Title: Superinfection immunity of mycobacteriophage L5: applications for genetic transduction
 A;Reference number: S30949; MUID:93211283; PMID:8459767
 A;Accession: S31005
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-47 <DON>
 A;Cross-references: UNIPROT:Q05273; EMBL:Z18946; NID:g15859; PIDN:CAA79436.1; PID:g1591
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992
 C;Genetics:
 A;Gene: 60

Query Match 25.8%; Score 25; DB 2; Length 47;
 Best Local Similarity 45.5%; Pred. No. 2.3e+03;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 IAFKIVSQEPA 19
 : || || : || || :
 DB 3 VSFKILGVEIA 13

RESULT 51

A30504
 Ig gamma heavy chain disease protein (RIV) - human
 C;Species: Homo sapiens (man)
 C;Date: 10-Nov-1988 #sequence_revision 03-Aug-1992 #text_change 16-Aug-1996
 C;Accession: A30504
 R;Guglielmi, P.; Bakhshi, A.; Cogne, M.; Seligmann, M.; Korsmeyer, S.J.
 J. Immunol. 141, 1762-1768, 1988
 A;Title: Multiple genomic defects result in an alternative RNA splice creating a human gamma heavy chain disease protein (RIV)
 A;Reference number: A30504; MUID:8315789; PMID:3137265
 A;Accession: A30504
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-48 <GUG>
 A;Note: the authors translated the codon GAG for residue 34 as Gly
 C;Keywords: immunoglobulin

Query Match 25.8%; Score 25; DB 2; Length 48;
 Best Local Similarity 35.3%; Pred. No. 2.3e+03;
 Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 PNHLNSKIAFKIVSQEP 18
 : || : || : || : ||
 DB 19 PRLVHSGVAVKRPQAE 35

RESULT 52

S72793
 cysteine synthase (EC 4.2.99.8) B - Mycobacterium leprae
 N;Alternate names: B1549_C3_238 protein; O-Acetylserine sulphydrolase B
 C;Species: Mycobacterium leprae
 C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C;Accession: S72793
 R;Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, November 1993
 A;Description: Mycobacterium leprae cosmid B1549.
 A;Reference number: S72582

A;Accession: S72793
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-49 <SMI>
C;Cross-references: UNIPROT:Q59529; EMBL:U00014; NID:g466903; PIDN:AAA50898.1; PID:g4669
C;Genetics:
A;Gene: cysM
C;Keywords: carbon-oxygen lyase

Query Match 25.8%; Score 25; DB 2; Length 49;
Best Local Similarity 71.4%; Pred. No. 2.4e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 KIVSQEP 18
|||:|
Db 15 KIVAAEP 21

RESULT 53

H97798
hypothetical protein RC0792 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: H97798
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: H97798
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <KUR>
A;Cross-references: UNIPROT:Q92HH9; GB:AE006914; PIDN:AAU0330.1; PID:gl5619890; GSPDB:G
C;Genetics:
A;Gene: RC0792

Query Match 25.8%; Score 25; DB 2; Length 50;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
|:|:|
Db 26 FEFIVQEP 33

RESULT 54

F95089
hypothetical protein SP0773 [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95089
R;Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
n, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95089
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-32 <KUR>
A;Cross-references: UNIPROT:Q97RN0; GB:AE005672; PIDN:AAK74911.1; PID:gl4972248; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0773

Query Match 25.3%; Score 24.5; DB 2; Length 32;
Best Local Similarity 42.9%; Pred. No. 1.8e+03;
Matches 6; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 3 NHLNSKIAFKIVSQ 16
|||:|

Db 19 NHIRMGV-FKIMFQ 31

RESULT 55

A39965
macrocin-O-methyltransferase (BC 2.1.1.-) - Streptomyces fradiae (fragment)
C;Species: Streptomyces fradiae
C;Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 09-Jul-2004
C;Accession: A39965
R;Fishman, S.E.; Cox, K.; Larson, J.L.; Reynolds, P.A.; Seno, E.T.; Yeh, W.K.; Van Fran
Proc. Natl. Acad. Sci. U.S.A. 84, 8248-8252, 1987
A;Title: Cloning genes for the biosynthesis of a macrolide antibiotic.
A;Reference number: A39965; MUID:88068571; PMID:3479787

A;Accession: A39965
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-34 <FIS>
A;Cross-references: UNIPROT:Q9S4D5; UNIPROT:Q54169; GB:J03008
C;Keywords: methyltransferase

Query Match 25.3%; Score 24.5; DB 2; Length 34;
Best Local Similarity 20.8%; Pred. No. 1.9e+03;
Matches 5; Conservative 6; Mismatches 6; Indels 7; Gaps 1;

QY 2 PNH-----LNSKIAFKIVSQEP 18
|:|:|:|
Db 3 PDEARDLYELLKKVSNVIYEDP 26

RESULT 56

T01741
hypothetical protein 2 - maize mitochondrion
C;Species: mitochondrion Zea mays (maize)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01741
R;Smith, A.G.; Pring, D.R.
Curr. Genet. 12, 617-623, 1987
A;Title: Nucleotide sequence and molecular characterization of a maize mitochondrial pl
A;Reference number: Z14414; MUID:89003167; PMID:2458851
A;Accession: T01741
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-38 <SMI>
A;Cross-references: UNIPROT:Q36273; EMBL:M36398; NID:g342649; PIDN:AAA70276.1; PID:g3426
C;Genetics:
A;Genome: mitochondrion
C;Keywords: mitochondrion

Query Match 25.3%; Score 24.5; DB 2; Length 38;
Best Local Similarity 36.8%; Pred. No. 2.2e+03;
Matches 7; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 PNLNSKIAP-KIVSQEPA 19
|:|:|:|
Db 20 PLHYRGLTFPRVRSREGA 38

RESULT 57

JT0514
IG heavy chain V-III region (JP4) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 23-Oct-1992 #sequence_revision 23-Oct-1992 #text_change 30-May-1997
C;Accession: JT0514
R;Anker, R.; Conley, M.E.; Pollok, B.A.
J. Exp. Med. 169, 2109-2119, 1989
A;Title: Clonal diversity in the B cell repertoire of patients with X-linked agammaglobu
A;Reference number: JT0511; MUID:89279157; PMID:2786547
A;Accession: JT0514
A;Molecule type: mRNA
A;Residues: 1-42 <ANK>
A;Note: the sequence shown here is one of eight productive V-D-J mu chain rearrangements
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

F;1-27/Domain: V region <VRE>
F;28-39/Domain: D region <DRE>
F;40-42/Domain: J region <JRE>

Query Match	25.3%;	Score	24.5;	DB 2;	Length	42;
Best Local Similarity	29.4%;	Pred. No.	2.4e+03;			
Matches	5;	Conservative	4;	Mismatches	7;	Indels
						1;
						Gaps
						1;

Qy 3 NHLNSKIAFKIVSQEPA 19
|| : : : |||
Db 6 NHFSLNLT-SVTGREPA 21

RESULT 58

I51287
calcium-binding protein, intestinal - chicken (fragment)
N:Alternate names: calbindin D9K; cholecalciferol; mammalian type CaBP; small CaBP
C/Species: Gallus gallus (Chicken)
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C/Accession: I51287
R/Zanillo, S.B.; Boland, R.L.; Norman, A.W.
Endocrinology 136, 2784-2787, 1995
A>Title: cDNA sequence identity of a vitamin D-dependent calcium-binding protein in the
A:Reference number: I51287; MUID:95269726; PMID:7750504

Query Match 25.3%; Score 24.5; DB 2; Length 50;
Best Local Similarity 43.8%; Pred. No. 2.9e+03;
Matches 7: Conservative 3: Mismatches 5: Indels

Qy 1 EPNHNSKIAFKIVSQ 16
:| | | | | | | |
Db 2 DPNOL-SKEELKLIO 16

RESULT 59

I77474
 myoglobin - Sericornis magnirostris (fragment)
 C:Species: Sericornis magnirostris
 C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
 C:Accession: I77474
 R:Slade, R.W.; Moritz, C.; Heideman, A.; Hale, P.T.
 Mol. Ecol. 2, 359-373, 1993
 A:Title: Rapid assessment of single-copy nuclear DNA variation in diverse species.
 A:Reference number: 157685; MUID:94214718; PMID:7909260
 A:Accession: I77474
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-25 <SLA>
 A:Cross-references: UNIPROT:Q91455; GB:L77493; NID:g310597; PIDN:AAA49619.1; PID:g310597
 C:Keywords: oxygen carrier

Query Match 24.7%; Score 24; DB 2; Length 25;
Best Local Similarity 23.1%; Pred. No. 1.7e+03;
Matches 3: Conservative 7: Mismatches 3: Indels

Qy 7 SKIAFKIVSQEPA 19
|::|:::|
Db 3 SEVIIKVIAEKHA 15

RESULT 60

hypothetical protein D-26 (MAT a2 5' region) - yeast (*Saccharomyces cerevisiae*)
A01640
C:Species: *Saccharomyces cerevisiae*
C:Date: 31-Dec-1990 #sequence revision 05-Sep-1996 #text change 08-Dec-2000

C:Accession: A01640
R:Astell, C.R.; Ahlstr
Cell 27, 15-23, 1981
A>Title: The sequence
A:Reference number: A9
A:Accession: A01640
A:Status: conceptual t
A:Molecule type: DNA
A:Residues: 1-26 <AST>
A:Cross-references: EM
C:Comment: This is the
C:Genetics:
A:Map position: 3R
C:Keywords: pseudogene

Query Match	24.7%	Score 24;	DB 4;	Length 26;
Best Local Similarity	37.5%	Pred. No. 1.7e+03;		
Matches	6;	Conservative	3;	Mismatches 7;
				Indels 0;
				Gaps 0;

QY 4 HLNSKIAFKIVSQEPA 19
||| : |::|
db 10 HLNSLLPPLYYSKKCA 25

RESIT.T 61

S72460
ribosomal protein S19 - curled-leaved tobacco chloroplast (fragment)
C:Species: chloroplast Nicotiana glauca (curled-leaved tobacco)
C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C/Accession: S72460
R:Goulding, S.E.; Olmstead, R.G.; Morden, C.W.; Wolfe, K.H.
Mol. Gen. Genet. 252, 195-206, 1996
A/Title: Bbb and flow of the chloroplast inverted repeat.
A/Reference number: S72459; MUID:96397499; PMID:8804393
A/Accession: S72460
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-28 <GU>
A/Cross-references: UNIPROT:Q36593; EMBL:Z71241; NID:gl279593; PIDN:CAA9494
A/Note: only a part of the nucleic acid sequence is shown
A/Note: the nucleotide sequence was submitted to the EMBL Data Library. Apr
C/Genetics:

Query Match	24.7%	Score 24;	DB 2;	Length 28;
Best Local Similarity	71.4%	Pred. No. 1.9e+03;		
Matches	5;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;

Qy 3 NHLNSKI 9
db 13 NHLKKI 19

RESULT 62

GCCB
 glucagon - Chinchilla brevicaudata
 C:Species: Chinchilla brevicaudata, Chinchilla lanigera brevicaudata
 C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
 C:Accession: A60413
 R:Eng, J.; Kleinman, W.A.; Chu, L.S.
 Peptides 11, 683-685, 1990
 A:Title: Purification of peptide hormones from chinchilla pancreas by chemical assay.
 A:Reference number: A60413; PMID:91045327; PMID:2235678
 A:Accession: A60413
 A:Molecule type: protein
 A:Residues: 1-29 <ENG>
 A:Cross-references: UNIPROT:P31297

C;Superfamily: glucagon
C;Keywords: carbohydrate metabolism; duplication; hormone; pancreas

Query Match 24.7%; Score 24; DB 1; Length 29;
Best Local Similarity 45.5%; Pred. No. 2e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HNSKIAFKIV 14
|||:|:|:|:
Db 13 HLDSEYAEFV 23

RESULT 63

D95224
hypothetical protein SP1921 [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: D95224
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: D95224
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-31 <KUP>
A;Cross-references: UNIPROT:Q9NU1; GB:A8005672; PIDN:AAK75989.1; PID:g14973424; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SPI921

Query Match 24.7%; Score 24; DB 2; Length 31;
Best Local Similarity 38.5%; Pred. No. 2.1e+03;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQE 17
|:|:|:|:
Db 3 LSEKITWDFNQE 15

RESULT 64

AD0740
hypothetical protein STY2078 [imported] - Salmonella enterica subsp. enterica serovar Ty
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AD0740
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AD0740
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-38 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05623.1; PID:g16503120; GSPDB:GN00176
C;Genetics:
A;Gene: STY2078

Query Match 24.7%; Score 24; DB 2; Length 38;
Best Local Similarity 57.4%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
|||:|:|:
Db 10 NHLRHKL 16

RESULT 65

S72464
ribosomal protein S19 - tomato chloroplast (fragment)
C;Species: Chloroplast lycopersicon esculentum (tomato)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: S72464
R;Goulding, S.E.; Olmstead, R.G.; Morden, C.W.; Wolfe, K.H.
Mol. Gen. Genet. 252, 195-206, 1996
A;Title: Ebb and flow of the chloroplast inverted repeat.
A;Reference number: S72459; MUID:96397499; PMID:8804393
A;Accession: S72464
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-40 <GOU>
A;Cross-references: UNIPROT:Q36781; EMBL:Z71246; NID:g1279722; PIDN:CAA94958.1; PID:g43;
A;Experimental source: cv. Moneymaker
A;Note: only a part of the nucleic acid sequence is shown
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1996
C;Genetics:
A;Gene: rps19
A;Genome: chloroplast
A;Start codon: GTG
C;Function:
A;Pathway: protein biosynthesis
C;Superfamily: ribosomal protein S19/S15
C;Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 24.7%; Score 24; DB 2; Length 40;
Best Local Similarity 71.4%; Pred. No. 2.8e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
|||:|:|:
Db 13 NHLHKKI 19

RESULT 66

S25728
beta-lactamase (EC 3.5.2.6) II - Citrobacter diversus (strain ULA27) (fragments)
C;Species: Citrobacter diversus
A;Variety: strain ULA27
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 03-May-1996
C;Accession: S25728
R;Franceschini, N.; Amicosante, G.; Perilli, M.; Maccarrone, M.; Oratore, A.; van Beeum
Biochem. J. 275, 629-633, 1991
A;Title: Proteolytic interconversion and N-terminal sequences of the Citrobacter divers
A;Reference number: S15418; MUID:91248093; PMID:2039443
A;Accession: S25728
A;Molecule type: protein
A;Residues: 1-38;39-41 <FRA>
C;Superfamily: beta-lactamase I
C;Keywords: antibiotic resistance; hydrolase

Query Match 24.7%; Score 24; DB 2; Length 41;
Best Local Similarity 40.8%; Pred. No. 2.9e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EPNHLSKIA 10
|||:|:|:
Db 4 EPTQVQKLA 13

RESULT 67

T18202
hypothetical protein - Bacillus sp. (fragment)
C;Species: Bacillus sp.
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18202
R;van Waasbergen, L.G.; Hildebrand, M.; Tebo, B.M.
J. Bacteriol. 178, 3517-3530, 1996
A;Title: Identification and characterization of a gene cluster involved in manganese oxi

A;Reference number: Z18825; MUID:96256605; PMID:8655549
 A;Accession: F18202
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-41 <VAN>
 A;Cross-references: UNIPROT:P70956; EMBL:U31081; NID:G942616; PID:G1502369; PIDN:AAB0648

Query Match 24.7%; Score 24; DB 2; Length 41;
 Best Local Similarity 36.4%; Pred. No. 2.9e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 NLSKIAPKIV 13
 ||| : : |||

Db 22 NHVITSLMFNI 32

RESULT 68
 AC1594

hypothetical protein lin1292 [imported] - Listeria innocua (strain Clip11262)

C;Species: Listeria innocua
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C;Accession: AC1594
 R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, R.; Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AC1594

A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-41 <GLA>

A;Cross-references: UNIPROT:Q925Y3; GB:AL592022; PIDN:CAC96523.1; PID:G16413765; GSPDB:G
 A;Experimental source: strain Clip11262
 C;Genetics:
 A;Gene: lin1292

Query Match 24.7%; Score 24; DB 2; Length 41;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 NLSKIAPKIV 14
 ||| : : |||

Db 10 LNGKTFKEV 19

RESULT 69
 A11645

protein gp22 (Bacteriophage A118) homolog lin1706 [imported] - Listeria innocua (strain

C;Species: Listeria innocua
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C;Accession: A11645
 R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, R.; Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: A11645

A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-41 <GLA>

A;Cross-references: UNIPROT:Q925Y3; GB:AL592022; PIDN:CAC96937.1; PID:G16414193; GSPDB:G
 A;Experimental source: strain Clip11262
 C;Genetics:
 A;Gene: lin1706

Query Match 24.7%; Score 24; DB 2; Length 41;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 5 NLSKIAPKIV 14
 ||| : : |||

Db 10 LNGKTFKEV 19

RESULT 70

S72462

ribosomal protein S19 - Nicotiana glauca chloroplast (fragment)
 C;Species: Chloroplast Nicotiana glauca
 C;Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
 C;Accession: S72462
 R;Goulding, S.E.; Olmstead, R.G.; Morden, C.W.; Wolfe, K.H.
 Mol. Gen. Genet. 252, 195-206, 1996
 A;Title: Ebb and flow of the chloroplast inverted repeat.
 A;Reference number: S72459; MUID:96397499; PMID:8804393
 A;Accession: S72462

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-44 <GOU>

A;Cross-references: UNIPROT:Q36592; EMBL:Z71235; NID:G1279585; PIDN:CAA94935.1; PID:943

A;Note: only a part of the nucleic acid sequence is shown

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1996

C;Genetics:

A;Gene: rps19

A;Genome: chloroplast

A;Start codon: GTG

C;Function:

A;Pathway: protein biosynthesis

C;Superfamily: ribosomal protein S19/S15

C;Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 24.7%; Score 24; DB 2; Length 44;
 Best Local Similarity 71.4%; Pred. No. 3.1e+03;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NLSKSI 9

Db 13 NLSKSI 19

RESULT 71

D49281

pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (isolate Pat24) (fragment

C;Species: simian T-cell lymphotropic virus type 1, STLV-1

C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004

C;Accession: D49281

R;Saksena, N.K.; Herve, V.; Durand, J.P.; Leguenno, B.; Diop, O.M.; Digoutte, J.P.; Mat

i Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Foiesz, B.J.

Virology 198, 297-310, 1994

A;Title: Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leuke

A;Reference number: A49281; MUID:94082462; PMID:8259665

A;Accession: D49281

A;Status: preliminary; translation not shown

A;Molecule type: DNA

A;Residues: 1-46 <SAK>

A;Cross-references: UNIPROT:Q88403; GB:L20353; NID:G431711; PIDN:AAAI6557.1; PID:G43171

C;Superfamily: pol polyprotein

Query Match 24.7%; Score 24; DB 2; Length 46;
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NLSLN 6

Db 37 NLSLN 40

RESULT 72

A49281

pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (isolate Bab34) (fragment

C;Species: simian T-cell lymphotropic virus type 1, STLV-1

C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
 C;Accession: A49281
 R;Saksena, N.K.; Herve, V.; Durand, J.P.; Leguenno, B.; Diop, O.M.; Digoutte, J.P.; Math
 ; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Polesz, B.J.
 Virology 198, 297-310, 1994
 A;Title: Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leukem
 A;Reference number: A49281; MUID:94082462; PMID:8259665
 A;Accession: A49281
 A;Status: preliminary; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-46 <SAK>
 A;Cross-references: UNIPROT:Q88390; GB:L20351; NID:g431689; PIDN:AAA16596.1; PID:g431690
 C;Superfamily: pol polyprotein

Query Match 24.7%; Score 24; DB 2; Length 46;
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NHLN 6
 ||||
 Db 37 NHLN 40

RESULT 73
 F49281
 pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (fragment)
 C;Species: simian T-cell lymphotropic virus type 1, STLV-1
 C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
 C;Accession: F49281; B49281
 R;Saksena, N.K.; Herve, V.; Durand, J.P.; Leguenno, B.; Diop, O.M.; Digoutte, J.P.; Math
 ; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Polesz, B.J.
 Virology 198, 297-310, 1994
 A;Title: Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leukem
 A;Reference number: A49281; MUID:94082462; PMID:8259665
 A;Accession: F49281
 A;Status: preliminary; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-46 <SAK>
 A;Cross-references: UNIPROT:Q88344; UNIPROT:Q88360; UNIPROT:Q88343; UNIPROT:Q88394; UNIP
 ROT:Q88331; UNIPROT:Q88406; GB:L20357; NID:g431721; PIDN:AAA16561.1; PID:g431722
 A;Experimental source: isolate SAB37P01
 A;Accession: B49281
 A;Status: preliminary; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-46 <SAK>
 A;Cross-references: GB:L20358; NID:g431709; PIDN:AAA16556.1; PID:g431710
 A;Experimental source: isolate SAB91037
 C;Superfamily: pol polyprotein

Query Match 24.7%; Score 24; DB 2; Length 46;
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NHLN 6
 ||||
 Db 37 NHLN 40

RESULT 74
 E49281
 pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (isolate Pat74) (fragment)
 C;Species: simian T-cell lymphotropic virus type 1, STLV-1
 C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
 C;Accession: B49281
 R;Saksena, N.K.; Herve, V.; Durand, J.P.; Leguenno, B.; Diop, O.M.; Digoutte, J.P.; Math
 ; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Polesz, B.J.
 Virology 198, 297-310, 1994
 A;Title: Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leukem
 A;Reference number: A49281; MUID:94082462; PMID:8259665
 A;Accession: B49281
 A;Status: preliminary; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-46 <SAK>

A;Cross-references: UNIPROT:Q88346; UNIPROT:Q88404; UNIPROT:Q88345; GB:L20354; NID:g431
 C;Superfamily: pol polyprotein

Query Match 24.7%; Score 24; DB 2; Length 46;
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NHLN 6
 ||||
 Db 37 NHLN 40

RESULT 75
 C49281
 pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (isolate Mona22) (fragment)
 C;Species: simian T-cell lymphotropic virus type 1, STLV-1
 C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
 C;Accession: C49281
 R;Saksena, N.K.; Herve, V.; Durand, J.P.; Leguenno, B.; Diop, O.M.; Digoutte, J.P.; Math
 ; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Polesz, B.J.
 Virology 198, 297-310, 1994
 A;Title: Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leukem
 A;Reference number: A49281; MUID:94082462; PMID:8259665
 A;Accession: C49281
 A;Status: preliminary; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-46 <SAK>
 A;Cross-references: UNIPROT:Q88399; GB:L20352; NID:g431705; PIDN:AAA16555.1; PID:g431707
 C;Superfamily: pol polyprotein

Query Match 24.7%; Score 24; DB 2; Length 46;
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NHLN 6
 ||||
 Db 37 NHLN 40

RESULT 76
 S14022
 hypothetical protein 7A - Chlamydomonas reinhardtii transposon
 C;Species: Chlamydomonas reinhardtii
 C;Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
 C;Accession: S14022
 R;Day, A.; Rochaix, J.D.
 Nucleic Acids Res. 19, 1259-1266, 1991
 A;Title: A transposon with an unusual LTR arrangement from Chlamydomonas reinhardtii con
 A;Reference number: S14018; MUID:91232906; PMID:1851555
 A;Accession: S14022
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-47 <DAY>
 A;Cross-references: UNIPROT:Q99196; EMBL:X56231; NID:g18229; PIDN:CAA39684.1; PID:g13607
 C;Genetics:
 A;Mobile element: transposon

Query Match 24.7%; Score 24; DB 2; Length 47;
 Best Local Similarity 23.5%; Pred. No. 3.4e+03;
 Matches 4; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 PNHLSKIAFKIVSQEP 18
 | : : : | : | : |
 Db 31 PNPFGHVCIHVTSSEP 47

RESULT 77
 H64544
 ribosomal protein l32 - Helicobacter pylori (strain 26695)
 C;Species: Helicobacter pylori
 C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
 C;Accession: H64544
 R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weldman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*. A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: H64544

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-48 <TOM>

A;Cross-references: UNIPROT:P56054; GB:AE000540; GB:AE000511; NID:g2313287; PIDN:AAD0727

Query Match 24.7%; Score 24; DB 2; Length 48;
Best Local Similarity 60.0%; Pred. No. 3.4e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNHLN 6
Db 38 PHHIN 42

RESULT 78
G71964
ribosomal protein L32 - *Helicobacter pylori* (strain J99)
A;Species: *Helicobacter pylori*
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: G71964
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*.
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: G71964
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-48 <ARN>
A;Cross-references: UNIPROT:Q9ZMN2; GB:AE001456; GB:AE001439; NID:g4154689; PIDN:AAD0575
A;Experimental source: strain J99
C;Genetics:
A;Gene: rpmF

Query Match 24.7%; Score 24; DB 2; Length 48;
Best Local Similarity 60.0%; Pred. No. 3.4e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNHLN 6
Db 38 PHHIN 42

RESULT 79
C72257
hypothetical protein - *Thermotoga maritima* (strain MSB8)
A;Species: *Thermotoga maritima*
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: C72257
R;Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickley Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M. Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of *Thermotoga maritima*.
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: C72257
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <ARN>
A;Cross-references: UNIPROT:Q9X1C2; GB:AE001793; GB:AE000512; NID:g4981963; PIDN:AAD3647
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TWI402

Query Match 24.7%; Score 24; DB 2; Length 50;
Best Local Similarity 46.2%; Pred. No. 3.6e+03; Mismatches 5; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Best Local Similarity 30.8%; Pred. No. 3.6e+03; Mismatches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAPKIVSQE 17
Db 24 LNVIMFERSIKKK 36

RESULT 80
S06495
penicillin-binding protein 6 - *Enterococcus hirae* (fragment)
N;Contains: peptidyl dipeptidase (EC 3.4.15.-)
C;Species: *Enterococcus hirae*
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S06495
R;el Kharroubi, A.; Piras, G.; Jacques, P.; Szabo, I.; van Beuemen, J.; Coyette, J.; Ghosh, S.; Biochem. J. 262, 457-462, 1989
A;Title: Active-site and membrane topology of the DD-peptidase/penicillin-binding protein 6 of *Enterococcus hirae*.
A;Reference number: S06495; MUID:90026308; PMID:2803261
A;Accession: S06495
A;Molecule type: Protein
A;Residues: 1-50 <ELK>
A;Cross-references: UNIPROT:Q7MOL9
C;Keywords: peptidyl dipeptide hydrolase
F;35/Active site: Ser (covalent penicillin-binding) #status predicted

Query Match 24.7%; Score 24; DB 2; Length 50;
Best Local Similarity 44.4%; Pred. No. 3.6e+03; Mismatches 2; Indels 0; Gaps 0;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LNSKIAPKIVSQE 17
Db 24 LNVIMFERSIKKK 36

RESULT 81
G82540
hypothetical protein XF2564 [imported] - *Xylella fastidiosa* (strain 9a5c)
C;Species: *Xylella fastidiosa*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82540
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82540
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <SIM>
A;Cross-references: UNIPROT:Q9PAF4; GB:AE004064; GB:AE003849; NID:g9107775; PIDN:AAF8533
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.P.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, as-Neto, E.; Docena, C.; El-Dorzy, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laichado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.F.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaka, A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2564

Query Match 24.7%; Score 24; DB 2; Length 50;
Best Local Similarity 46.2%; Pred. No. 3.6e+03; Mismatches 5; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVS 15
 |||: | :||
 Db 34 NHLHLIFNHLVS 46

RESULT 82

S21694
 A:Title: thymosin beta-12 - Japanese seaperch
 C:Species: lateolabrax japonicus (Japanese seaperch)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S21694
 R:Low, T.L.K.; Liu, D.T.; Jou, J.
 Arch. Biochem. Biophys. 293, 32-39, 1992
 A:Title: Primary structure of thymosin beta(12), a new member of the beta-thymosin family
 A:Reference number: S21694; MUID:92117698; PMID:1731637
 A:Accession: S21694
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-43 <LOW>
 A:Cross-references: UNIPROT:P33248
 C:Superfamily: thymosin beta

Query Match 24.2%; Score 23.5; DB 2; Length 43;
 Best Local Similarity 42.1%; Pred. No. 3.7e+03;
 Matches 8; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 1 EPNHLNSKIAFKIVSQEPA 19
 |||: | :||
 Db 24 EKNPLPSK--ETTEQEKX 39

RESULT 83

S58644
 A:Title: tubulin beta chain - nematode (Haemonchus contortus) (isolate resistant South Africa) (5
 C:Species: Haemonchus contortus
 A:Variety: isolate resistant South Africa
 C:Date: 29-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
 R:Kwa, M.S.G.; Kooyman, F.N.J.; Boersma, J.H.; Roos, M.H.
 Biochem. Biophys. Res. Commun. 191, 413-419, 1993
 A:Title: Effect of selection for benzimidazole resistance in Haemonchus contortus on be
 A:Reference number: S58644; MUID:92123272; PMID:8096381
 A:Accession: S58644
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-48 <KWA>
 A:Cross-references: UNIPROT:Q08744; EMBL:X67488; NID:G396694; PIDN:CAA7828.1; PID:G3966
 A:Experimental source: isolate resistant South Africa; L3-larvae
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
 C:Genetics:
 A:Gene: RSA-2
 C:Superfamily: tubulin

Query Match 24.2%; Score 23.5; DB 2; Length 48;
 Best Local Similarity 33.3%; Pred. No. 4.2e+03;
 Matches 5; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 3 NHLNSKIAFKIVSQE 17
 |||: | :||
 Db 14 NQIGSKF-WVISDE 27

RESULT 84

G86209
 A:Title: protein F22G5.12 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: G86209
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G86209
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-48 <STO>
 A:Cross-references: UNIPROT:Q9LNX3; GB:AE005172; NID:G878542; PIDN:AAF79550.1; GSPDB:G
 C:Genetics:
 A:Gene: F22G5.12
 A:Map position: 1

Query Match 24.2%; Score 23.5; DB 2; Length 48;
 Best Local Similarity 41.7%; Pred. No. 4.2e+03;
 Matches 5; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 3 NHLNSKIAFKIV 14
 |||: | :||
 Db 7 NHAN---AFNVI 15

RESULT 85

S07700
 A:Title: T-cell receptor alpha chain J segment (34) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 23-Jul-1999
 C:Accession: S07700
 R:Yague, J.; Blackman, M.; Born, W.; Marrack, P.; Kappler, J.; Palmer, E.
 Nucleic Acids Res. 16, 1135-11364, 1988
 A:Title: The structure of V-alpha and J-alpha segments in the mouse.
 A:Reference number: S06466; MUID:89083566; PMID:2849763
 A:Accession: S07700
 A:Molecule type: mRNA
 A:Residues: 1-23 <YAG>
 A:Cross-references: EMBL:M38676; NID:G201210; PIDN:AAA40194.1; PID:G201211
 A:Experimental source: strain C57BL/Ka
 C:Genetics:
 A:Map position: 14
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: glycoprotein; heterodimer; T-cell receptor
 F:1-21/Domain: J segment <USE>
 F:22-23/Domain: C region (fragment) <CRE>

Query Match 23.7%; Score 23; DB 2; Length 23;
 Best Local Similarity 36.4%; Pred. No. 2.3e+03;
 Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHLNSKIAFP 11
 |||: | :||
 Db 1 EPTQMPYKVF 11

RESULT 86

S39391
 A:Title: calpain II heavy chain - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 08-Jun-1994 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
 C:Accession: S39391
 R:Crawford, C.; Brown, N.R.; Willis, A.C.
 Biochem. J. 296, 135-142, 1993
 A:Title: Studies of the active site of m-calpain and the interaction with calpastatin.
 A:Reference number: S39391; MUID:94071815; PMID:8250833
 A:Accession: S39391
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-25 <CRA>
 A:Cross-references: UNIPROT:Q7M386
 Query Match 23.7%; Score 23; DB 2; Length 25;
 Best Local Similarity 37.5%; Pred. No. 2.5e+03;

Matches 6; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

Qy 4 HLNS--KIAFKIVSQE 17
:||||| : : : :
Db 8 YLNSKKADYQWDDE 23

RESULT 87

I37301
MHC class II histocompatibility antigen HLA-DR beta-3 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C:Accession: I37301; I37302
R:Louis, P.; Eliaou, J.F.; Kerlan-Candon, S.; Pinet, V.; Vincent, R.; Clot, J.
Immunogenetics 38, 21-26, 1993
A:Title: Polymorphism in the regulatory region of HLA-DRB genes correlating with haplotypes
A:Reference number: I37300; MUID:93216303; PMID:8462990
A:Accession: I37301
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-29 <RES>
A:Cross-references: UNIPROT:Q9UBW6; EMBL:X65558; NID:g296268; PIDN:CAA46528.1; PID:g296262
A:Note: this allele is designated DRB3*0101
A:Accession: I37302
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-29 <RES>
A:Cross-references: EMBL:X65559; NID:g296270; PIDN:CAA46529.1; PID:g296271
A:Note: this allele is designated DRB3*0201
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 23.7%; Score 23; DB 2; Length 29;
Best Local Similarity 57.1%; Pred. No. 2.9e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LNSKIAF 11
: : : : :
Db 22 LSSRLAF 28

RESULT 88

S11617
ribosomal protein S24-eR [validated] - Halobacterium salinarum (fragment)
N:Alternate names: ribosomal protein HS17
C:Species: Halobacterium salinarum
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S11617
R:Yaguchi, M.; Visentin, L.P.; Zuker, M.; Matheson, A.T.; Roy, C.; Strom, A.R.
Zbl. Bakt. Hyg. I. Abt. Orig. C 3, 200-208, 1982
A:Title: Amino-terminal sequences of ribosomal proteins from the 30S subunit of archaea
A:Reference number: S11609
A:Accession: S11617
A:Molecule type: protein
A:Residues: 1-30 <YAG>
A:Cross-references: UNIPROT:OTM550
A:Note: the protein is designated as ribosomal protein HS17
A:Note: the source is designated as Halobacterium cutirubrum
C:Superfamily: rat ribosomal protein S24
C:Keywords: protein biosynthesis; ribosome

Query Match 23.7%; Score 23; DB 2; Length 30;
Best Local Similarity 36.8%; Pred. No. 3e+03; Mismatches 7; Indels 7; Gaps 1;

Qy 1 EPNHL--NSKIAFKIVSQE 17
: : : : :
Db 9 EDNPLIHRDVGFKIVHND 27

RESULT 89

G82528
hypothetical protein XF2676 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82528

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below

A:Accession: G82528
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-30 <SIM>
A:Cross-references: UNIPROT:Q9PA44; GB:AE004073; GB:AE003849; NID:g9107904; PIDN:AAF854
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, as-Neto, E.; Docena, C.; El-Dozry, H.; Paciniani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Laig, chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2676

Query Match 23.7%; Score 23; DB 2; Length 30;
Best Local Similarity 66.7%; Pred. No. 3e+03; Mismatches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KIAFKI 13
: : : : :
Db 10 EIAFKV 15

RESULT 90

E95866
probable pyrroloquinoline quinone synthesis protein A [imported] - Sinorhizobium meliloti
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: E95866
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna, Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-31 <KUR>
A:Cross-references: UNIPROT:Q9EXV2; GB:AL591985; PIDN:CAC48597.1; PID:g15140069; GSPDB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F. L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.; A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: pqqA; SMb20204
A:Genome: plasmid

Query Match 23.7%; Score 23; DB 2; Length 31;
Best Local Similarity 26.3%; Pred. No. 3.2e+03; Mismatches 5; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSQEPA 19

Db 5 KPKFIEVSCAMEITRYADA 23

RESULT 91

C82287
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C;Accession: C82287
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F. I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A;Reference number: A82035; MUID:20406833; PMID:10952301
 A;Accession: C82287
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-33 <HEI>
 A;Cross-references: UNIPROT:Q9KU01; GB:AE004159; GB:AE003852; NID:g9655167; PIDN:AAF9389
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VC0729
 A;Map position: 1

Query Match 23.7%; Score 23; DB 2; Length 33;
 Best Local Similarity 75.0%; Pred. No. 3.4e+03;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNHL 5
 Db 23 PNHI 26

RESULT 92

F82443
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C;Accession: F82443
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F. I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A;Reference number: A82035; MUID:20406833; PMID:10952301
 A;Accession: F82443
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-35 <HEI>
 A;Cross-references: UNIPROT:Q9KM12; GB:AE004388; GB:AE003853; NID:g9657979; PIDN:AAF9647
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VCA0577
 A;Map position: 2

Query Match 23.7%; Score 23; DB 2; Length 35;
 Best Local Similarity 55.6%; Pred. No. 3.6e+03;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NHLNSKIAPF 11
 Db 2 NYLPSAISF 10

RESULT 93

H70251
 C;Species: Borrelia burgdorferi (Lyme disease spirochete)
 C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 C;Accession: H70251
 R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whitte

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997
 A;Authors: Smith, H.O.; Venter, J.C.
 A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
 A;Reference number: A70100; MUID:98065943; PMID:9403685
 A;Accession: H70251
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-36 <KLE>
 A;Cross-references: UNIPROT:O50810; GB:AE000788; NID:g2690123; PIDN:AAC66152.1; PID:g26
 A;Experimental source: strain B31
 C;Genetics:
 A;Genome: plasmid

Query Match 23.7%; Score 23; DB 2; Length 36;
 Best Local Similarity 45.5%; Pred. No. 3.7e+03;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 NSKIAPKIVSQ 16
 Db 10 NPKIPLKISR 20

RESULT 94

S70801
 C;Species: Vibrio cholerae
 C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
 C;Accession: S70801
 R;Camilli, A.; Mekalanos, J.J.
 Mol. Microbiol. 18, 671-683, 1995
 A;Title: Use of recombinase gene fusions to identify Vibrio cholerae genes induced durin
 A;Reference number: S70798; MUID:96414469; PMID:8817490
 A;Accession: S70801
 A;Status: preliminary; nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 1-37 <CAM>
 A;Cross-references: UNIPROT:Q56601; EMBL:U25711; NID:g1165179; PIDN:AAC43553.1; PID:g11
 C;Superfamily: Denitrication system component NapC/NiRT (membrane-bound tetraheme cyto
 C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein
 F;12,15/Binding site: heme (Cys) (covalent) #status predicted
 F;16/Binding site: heme iron (His) (axial ligand) #status predicted

Query Match 23.7%; Score 23; DB 2; Length 37;
 Best Local Similarity 41.7%; Pred. No. 3.8e+03;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PNHLNSKIAPKI 13
 Db 18 PHEWTAKIARKM 29

RESULT 95

A23617
 C;Species: Lupinus angustifolius (narrow-leaved blue lupine)
 C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
 C;Accession: A23617
 R;Lilley, G.G.; Inglis, A.S.
 FEBS Lett. 195, 235-241, 1986
 A;Title: Amino acid sequence of conglutinin delta, a sulfur-rich seed protein of Lupinus a
 A;Reference number: A91358
 A;Accession: A23617
 A;Molecule type: protein
 A;Residues: 1-37 <LIL>
 A;Cross-references: UNIPROT:P09930
 C;Superfamily: soybean 2S albumin

Query Match 23.7%; Score 23; DB 2; Length 37;
 Best Local Similarity 42.9%; Pred. No. 3.8e+03;
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
 ||: : ||
 Db 22 NHIDQRI 28

RESULT 96
 T06924
 Photosystem II protein psbI - Cyanophora paradoxa cyanelle
 C:Species: cyanelle Cyanophora paradoxa
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
 C:Accession: T06924
 R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.
 submitted to the EMBL Data Library, July 1995
 A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
 A:Reference number: Z15840
 A:Accession: T06924
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-38 <STI>
 A:Cross-references: UNIPROT:P48106; EMBL:U0821; NID:gl016083; PIDN:AAA81267.1; PID:gl016083
 A:Experimental source: strain Pringsheim LB555
 C:Genetics:
 A:Gene: psbI
 A:Genome: cyanelle
 C:Superfamily: photosystem II protein psbI
 C:Keywords: cyanelle; membrane-associated complex; photosynthesis; photosystem II; thylakoid

Query Match 23.7%; Score 23; DB 2; Length 38;
 Best Local Similarity 38.5%; Pred. No. 4e+03;
 Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 7 SKIAPKIVSQEPA 19
 | | : : ||
 Db 17 SLTFGFLSDNPA 29

RESULT 97
 S72459
 Ribosomal protein S19 - Digitalis purpurea chloroplast (fragment)
 C:Species: chloroplast Digitalis purpurea
 C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
 C:Accession: S72459
 R:Goulding, S.E.; Olmstead, R.G.; Morden, C.W.; Wolfe, K.H.
 Mol. Gen. Genet. 252, 195-206, 1996
 A:Title: Ebb and flow of the chloroplast inverted repeat.
 A:Reference number: S72459; MUID:96397499; PMID:8804393
 A:Accession: S72459
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-39 <GOU>
 A:Cross-references: UNIPROT:Q33418; EMBL:Z71252; NID:gl279387; PIDN:CAA94970.1; PID:gl279387
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1996
 C:Genetics:
 A:Gene: rps19
 A:Genome: chloroplast
 A:Start codon: GTG
 C:Function:
 A:Pathway: protein biosynthesis
 C:Superfamily: ribosomal protein S19/S15
 C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 23.7%; Score 23; DB 2; Length 39;
 Best Local Similarity 71.4%; Pred. No. 4.1e+03;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
 ||: : ||
 Db 13 NHLRKI 19

RESULT 98
 C97513

hypothetical protein AGR_C_2336 [imported] - Agrobacterium tumefaciens (strain C58, Cer
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C:Accession: C97513
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: C97513
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-39 <KUR>
 A:Cross-references: UNIPROT:Q8U5C6; GB:AE007869; PIDN:AAK87060.1; PID:gl5156314; GSPDB:
 C:Genetics:
 A:Gene: AGR_C_2336
 A:Map position: circular chromosome

Query Match 23.7%; Score 23; DB 2; Length 39;
 Best Local Similarity 50.0%; Pred. No. 4.1e+03;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PNHLNSKI 9
 | | : ||
 Db 2 PRHASCKI 9

RESULT 99
 I49405
 beta-galactoside-binding lectin - western wild mouse (fragment)
 C:Species: Mus spretus (western wild mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C:Accession: I49405
 R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.
 Mamm. Genome 5, 349-355, 1994
 A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
 A:Reference number: I48934; MUID:94319082; PMID:8043949
 A:Accession: I49405
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-40 <RES>
 A:Cross-references: UNIPROT:Q62529; EMBL:U05693; NID:g497012; PIDN:AAB60464.1; PID:g497012
 C:Superfamily: beta-galactoside-binding lectin

Query Match 23.7%; Score 23; DB 2; Length 40;
 Best Local Similarity 80.0%; Pred. No. 4.2e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNHLN 6
 | | | |
 Db 15 ENRLN 19

RESULT 100
 E83742
 hypothetical protein BH0741 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C:Accession: E83742
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: E83742
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-41 <STO>
 A:Cross-references: UNIPROT:Q9KEV8; GB:AP001509; GB:BA000004; NID:gl0173176; PIDN:BA00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH0741

Query Match 23.7%; Score 23; DB 2; Length 41;

Best Local Similarity 50.0%; Pred. No. 4.3e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSK 8
|||.|||
Db 7 EPHSHLK 14

RESULT 101
A97948
hypothetical protein ABC-NBD-truncation [imported] - Streptococcus pneumoniae (strain R6)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: A97948
R/Hoskins, J.A.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
y, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; PMID:21429245; PMID:11544234
A/Accession: A97948
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-41 <KUR>
A/Cross-references: UNIPROT:Q8DQL9; GB:AE007317; PIDN:AAK99413.1; PID:g15458192; GSPDB:C
C/Genetics:
A/Gene: ABC-NBD-truncation

Query Match 23.7%; Score 23; DB 2; Length 41;
Best Local Similarity 38.5%; Pred. No. 4.3e+03;
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKI 13
|||.|||
Db 17 EVSHLSKSFQDKI 29

RESULT 102
I37286
olfactory receptor I12 - rat (fragment)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004
C/Accession: I37286
R/Buck, L.; Axel, R.
Cell 65, 175-187, 1991
A/Title: A novel multigene family may encode odorant receptors: a molecular basis for od
A/Reference number: A23701; PMID:91191556; PMID:1840504
A/Accession: I37286
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-44 <BUC>
A/Cross-references: UNIPROT:Q04058; GB:M64390; NID:g205841; PIDN:AAA1753.1; PID:g205842
C/Supfamily: Olfactory receptor OR14
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 23.7%; Score 23; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 4.7e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 FKIVS 15
|||||
Db 27 FKIVS 31

RESULT 103
E37286
olfactory receptor FI3 - rat (fragment)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004
C/Accession: E37286
R/Buck, L.; Axel, R.
Cell 65, 175-187, 1991

A/Title: A novel multigene family may encode odorant receptors: a molecular basis for c
A/Reference number: A23701; PMID:91191556; PMID:1840504
A/Accession: E37286
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-44 <BUC>
A/Cross-references: UNIPROT:Q04054; GB:M64382; NID:g205825; PIDN:AAA1745.1; PID:g20582
C/Supfamily: olfactory receptor OR14
C/Keywords: G protein-coupled receptor; membrane protein

Query Match 23.7%; Score 23; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 4.7e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 FKIVS 15
|||||
Db 27 FKIVS 31

RESULT 104
E85938
hypothetical protein Z4177 [imported] - Escherichia coli (strain O157:H7, substrain EDL
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: E85938
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; PMID:21074935; PMID:11206551
A/Accession: E85938
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-44 <STO>
A/Cross-references: UNIPROT:Q8X3M1; GB:AB005174; NID:g12517352; PIDN:AAG57969.1; GSPDB:
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: Z4177

Query Match 23.7%; Score 23; DB 2; Length 44;
Best Local Similarity 57.1%; Pred. No. 4.7e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 KIVSQEP 18
|||.|||
Db 11 KIIXPEP 17

RESULT 105
AG2286
Cytochrome b559 beta chain [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AG2286
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807; PMID:21595285; PMID:11759840
A/Accession: AG2286
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-45 <KUR>
A/Cross-references: UNIPROT:Q8YQI1; GB:BA000019; PIDN:BA075545.1; PID:g17132980; GSPDB:C
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: psbF
C/Supfamily: cytochrome b559, component F

Query Match 23.7%; Score 23; DB 2; Length 45;
Best Local Similarity 27.3%; Pred. No. 4.8e+03;
Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 NHLNSKIAPKI 13
|:|:|:|:|
Db 5 NNINQVVTYPI 15

RESULT 106

A61295
creatine kinase (EC 2.7.3.2) chain M2 - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 09-Jul-2004
C:Accession: A61295
R:Takasawa, T.; Onodera, M.; Shiokawa, H.
J. Biochem. 93, 389-395, 1983
A:Title: Properties of three creatine kinases MM from porcine skeletal muscle.
A:Reference number: A61295; MUID:83186110; PMID:6404898
A:Accession: A61295
A:Molecule type: protein
A:Residues: 1-45 <TAK>
A:Cross-references: UNIPROT:Q7M337
C:Superfamily: creatine kinase; creatine kinase repeat homology
C:Keywords: heterodimer; homodimer; phosphotransferase

Query Match 23.7%; Score 23; DB 2; Length 45;
Best Local Similarity 33.3%; Pred. No. 4.8e+03;
Matches 5; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 4 HLNSKIAPKIVSQEP 18
|:|:|:|:|
Db 6 HNKYKLNFKAEIEYP 20

RESULT 107

B61295
creatine kinase (EC 2.7.3.2) chain M1 - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 17-Jul-1994 #sequence_revision 22-Jul-1994 #text_change 09-Jul-2004
C:Accession: B61295
R:Takasawa, T.; Onodera, M.; Shiokawa, H.
J. Biochem. 93, 389-395, 1983
A:Title: Properties of three creatine kinases MM from porcine skeletal muscle.
A:Reference number: A61295; MUID:83186110; PMID:6404898
A:Accession: B61295
A:Molecule type: protein
A:Residues: 1-45 <TAK>
A:Cross-references: UNIPROT:Q7M336
C:Superfamily: creatine kinase; creatine kinase repeat homology
C:Keywords: heterodimer; homodimer; phosphotransferase

Query Match 23.7%; Score 23; DB 2; Length 45;
Best Local Similarity 33.3%; Pred. No. 4.8e+03;
Matches 5; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 4 HLNSKIAPKIVSQEP 18
|:|:|:|:|
Db 6 HNKYKLNFKAEIEYP 20

RESULT 108

H83816
hypothetical protein BH1336 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: H83816
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: H83816
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-45 <STO>

A:Cross-references: UNIPROT:Q9KDB1; GB:AF001511; GB:BA000004; NID:g10173727; PIDN:BA0505

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH1336

Query Match 23.7%; Score 23; DB 2; Length 45;
Best Local Similarity 30.0%; Pred. No. 4.8e+03;
Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NHLNSKIAPK 12
|:|:|:|:|
Db 2 NHFEKDVQYK 11

RESULT 109

T07459
hypothetical protein 46b - Japanese black pine chloroplast
C:Species: chloroplast Pinus thunbergiana (Japanese black pine)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C:Accession: T07459
R:Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiyura, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A:Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genom
A:Reference number: Z16030; MUID:95024047; PMID:7937893

A:Accession: T07459
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-46 <WAK>
A:Cross-references: UNIPROT:Q32940; EMBL:D17510; NID:g529643; PIDN:BA04337.1; PID:g126
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 23.7%; Score 23; DB 2; Length 46;
Best Local Similarity 75.0%; Pred. No. 4.9e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPNH 4
|:|:|:|
Db 9 DPNH 12

RESULT 110

B36626
osteogenic protein 16K chain - bovine (fragments)
C:Species: Bos primigenius taurus (cattle)
C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 09-Jul-2004
C:Accession: B36626
R:SamPATH, T.K.; Coughlin, J.E.; Whetstone, R.M.; Banach, D.; Corbett, C.; Ridge, R.J.
J. Biol. Chem. 265, 13198-13205, 1990
A:Title: Bovine osteogenic protein is composed of dimers of OP-1 and BMP-2A, two member
A:Reference number: A36626; MUID:90330665; PMID:2376592

A:Accession: B36626
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-47 <SAM>
A:Cross-references: UNIPROT:Q7M301
C:Superfamily: inhibin

Query Match 23.7%; Score 23; DB 2; Length 47;
Best Local Similarity 66.7%; Pred. No. 5e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNHLNS 7
|:|:|:|
Db 19 PDLXS 24

RESULT 111

T37087
probable IS element ATP-binding protein truncated [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
C:Accession: T37087

RiSaunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M. submitted to the EMBL Data Library, August 1999
A:Reference number: Z21588
A:Accession: T37087
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-47 <SAU>
A:Cross-references: UNIPROT:Q8RSN1; EMBL:AL109950; PIDN:CAB52938.1; GSPDB:GN00070; SCOEED
A:Experimental source: strain A3(2)
A>Note: an incorrect termination codon was used
C:Genetics:
A:Gene: SCOEEDB-SCJ4.04c
C:Superfamily: DNA replication protein dnaC

Query Match 23.7%; Score 23; DB 2; Length 47;
Best Local Similarity 20.0%; Pred. No. 5e+03;
Matches 2; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 7 SKIAPKIVSQ 16
::: ||:::
Db 17 ARLVFOVSK 26

RESULT 112
E81833
hypothetical protein NMA2034 [imported] - Neisseria meningitidis (strain Z2491 serogroup C)
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: E81833
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: E81833
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-47 <PAR>
A:Cross-references: UNIPROT:Q9JTL6; GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB8525
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA2034

Query Match 23.7%; Score 23; DB 2; Length 47;
Best Local Similarity 40.0%; Pred. No. 5e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 8 KIAFKIVSQE 17
|: |||: |
Db 3 KVTMKIVTDK 12

RESULT 113
A24941
storage protein - flesh fly (Sarcophaga peregrina) (fragment)
C:Species: Sarcophaga peregrina
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C:Accession: A24941
R:Matsumoto, N.; Nakanishi, Y.; Natori, S.
Nucleic Acids Res. 14, 2685-2698, 1986
A:Title: Homologies of nucleotide sequences in the 5'-end regions of two developmentally
A:Reference number: A24941; MUID:86176758; PMID:3960729
A:Accession: A24941
A:Molecule type: DNA
A:Residues: 1-48 <MAT>
A:Cross-references: UNIPROT:Q26659; GB:M31352; NID:g161270; PIDN:AAA29886.1; PID:g161271
C:Superfamily: arylphorin

Query Match 23.7%; Score 23; DB 2; Length 48;
Best Local Similarity 50.0%; Pred. No. 5.1e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAPKIV 14

Db 33 LKQKFLFEIV 42

RESULT 114

F70247

hypothetical protein BBJ21 - Lyme disease spirochete plasmid J/lp38
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: F70247

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-48 <KLE>

A:Cross-references: UNIPROT:O50776; GB:AE000787; NID:g2690175; PIDN:AAC66117.1; PID:g26

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

Query Match

Best Local Similarity 23.7%; Score 23; DB 2; Length 48;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 9 IAFKIVSQEP 18

|: |||: |

Db 10 ISFKYVLLKP 19

RESULT 115

F82224

hypothetical protein VC1243 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: F82224

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, L.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: F82224

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-48 <HEI>

A:Cross-references: UNIPROT:Q9KSL3; GB:AE004203; GB:AE003852; NID:g9655716; PIDN:AAF944

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1243

A:Map position: 1

Query Match 23.7%; Score 23; DB 2; Length 48;

Best Local Similarity 66.7%; Pred. No. 5.1e+03;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NHIKSK 8

|: |||: |

Db 27 SHLNK 32

RESULT 116

B69173

ribosomal protein L40 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: B69173

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
ki, S.; Church, G.W.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcb
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69173
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-48 <MTH>
A:Cross-references: UNIPROT:O26653; GB:AE000838; GB:AE000666; NID:g2621625; PIDN:AAB8505
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH553

Query Match 23.7%; Score 23; DB 2; Length 48;
Best Local Similarity 38.5%; Pred. No. 5.1e+03;
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQP 18
Db 13 NIKICLKNARNP 25

RESULT 117
G69538
conserved hypothetical protein AF2311 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Gladak, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: G69538
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-49 <KLE>
A:Cross-references: UNIPROT:O27973; GB:AE000944; GB:AE000782; NID:g2689267; PIDN:AAB8894

Query Match 23.7%; Score 23; DB 2; Length 49;
Best Local Similarity 41.7%; Pred. No. 5.2e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQ 16
Db 19 LNSKEVSKIIE 30

RESULT 118
C41662
Probable pheromone production regulation protein 3 - Enterococcus faecalis plasmid pCF10
C:Species: Enterococcus faecalis
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Jun-1993
C:Accession: C41662
R:Kao, S.M.; Olmsted, S.B.; Viksins, A.S.; Gallo, J.C.; Dunny, G.M.
J. Bacteriol. 173, 7650-7664, 1991
A:Title: Molecular and genetic analysis of a region of plasmid pCF10 containing positive
terococcus faecalis.
A:Reference number: A41662; MUID:92041679; PMID:1938961
A:Accession: C41662
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-50 <KAO>
A:Cross-references: GB:M64978
C:Genetics:
A:Genome: plasmid

Query Match 23.7%; Score 23; DB 2; Length 50;
Best Local Similarity 41.7%; Pred. No. 5.4e+03;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 NMLNSKIAFKIV 14
Db 28 NNLNIYILIEKII 39

RESULT 119
F82409
hypothetical protein VCA0842 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82409
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
I. R.K.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: F82409
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-50 <HEL>
A:Cross-references: UNIPROT:Q9KUA4; GB:AE004412; GB:AE003853; NID:g9658269; PIDN:AAF967
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0842
A:Map position: 2

Query Match 23.7%; Score 23; DB 2; Length 50;
Best Local Similarity 30.8%; Pred. No. 5.4e+03;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 7 SKIAFKIVSQEPA 19
Db 10 TRLSIPIVMNSGPA 22

RESULT 120
D31753
Hypothetical Oct-2 protein (clone pass-3) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: D31753
R:Clerc, R.G.; Corcoran, L.M.; LeBowitz, J.H.; Baltimore, D.; Sharp, P.A.
Genes Dev. 2, 1570-1581, 1988
A:Title: The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains
A:Reference number: A91621; MUID:89107992; PMID:3265124
A:Accession: D31753
A:Molecule type: mRNA
A:Residues: 1-50 <CLE>
A:Cross-references: UNIPROT:Q16638; GB:M36772; NID:g189365; PIDN:AAA36391.1; PID:g189365

Query Match 23.7%; Score 23; DB 2; Length 50;
Best Local Similarity 30.0%; Pred. No. 5.4e+03;
Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
Db 5 LGRLVRPEP 14

RESULT 121
G82817
hypothetical protein XF0332 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82817
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below

A;Accession: G82817
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-36 <SIM>
 A;Cross-references: UNIPROT:Q9PGH0; GB:AE003886; GB:AE003849; NID:G9105157; PIDN:AAF8314
 A;Experimental source: strain 945C
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328
 A;Contents: annotation
 C;Genetics:
 A;Gene: XF0332

Query Match 23.2%; Score 22.5; DB 2; Length 36;
 Best Local Similarity 38.5%; Pred. No. 4.5e+03;
 Matches 5; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Qy 2 PNHLNS---KIAF 11
 | | | | |
 Db 12 PTHRSYFVKVRF 24

RESULT 122
 F82294
 hypothetical protein VC0670 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C;Accession: F82294
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A;Reference number: A82035; MUID:20406833; PMID:10952301
 A;Accession: F82294
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-48 <HEI>
 A;Cross-references: UNIPROT:Q9KU54; GB:AE004153; GB:AE003852; NID:G9655103; PIDN:AAF9383
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VC0670
 A;Map position: 1

Query Match 23.2%; Score 22.5; DB 2; Length 48;
 Best Local Similarity 75.0%; Pred. No. 6.2e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 3 NHLSKIA 10
 | | | | |
 Db 31 NH-NSKFA 37

RESULT 123
 S26130
 outer membrane protein class 1 - Neisseria meningitidis (fragment)
 C;Species: Neisseria meningitidis
 C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
 C;Accession: S26130
 R;McGuinness, B.T.
 submitted to the EMBL Data Library, August 1992
 A;Reference number: S26099
 A;Accession: S26130

A;Molecule type: DNA
 A;Residues: 1-50 <MCG>
 A;Cross-references: UNIPROT:Q51218; EMBL:Z14288; NID:G45139; PIDN:CAA78655.1; PID:G4514
 C;Genetics:
 A;Gene: porA
 C;Superfamily: outer membrane protein class 1
 C;Keywords: membrane protein

Query Match 23.2%; Score 22.5; DB 2; Length 50;
 Best Local Similarity 43.8%; Pred. No. 6.5e+03;
 Matches 7; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

Qy 6 NSKTAF---KIVSOEP 18
 | | | | |
 Db 7 NKSAYTPAHFVQPPP 22

RESULT 124
 S00774
 kinase-related transforming protein (abl) (EC 2.7.1.-) type II - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 18-Jun-1999
 C;Accession: S00774
 R;Bernards, A.; Paskind, M.; Baltimore, D.
 Oncogene 2, 297-304, 1988
 A;Title: Four murine c-abl mRNAs arise by usage of two transcriptional promoters and al;
 A;Reference number: S00771; MUID:88202920; PMID:3283651
 A;Accession: S00774
 A;Molecule type: DNA
 A;Residues: 1-20 <BER>
 A;Cross-references: EMBL:X07540; NID:G49835; PIDN:CAA30413.1; PID:G49836
 A;Note: the authors translated the codon GAC for residue 12 as His
 C;Genetics:
 A;Gene: abl
 C;Superfamily: human protein-tyrosine kinase abl; protein kinase homology; SH2 homology
 C;Keywords: alternative splicing; ATP; phosphotransferase

Query Match 22.7%; Score 22; DB 2; Length 20;
 Best Local Similarity 44.4%; Pred. No. 2.9e+03;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 9 IAFKIVSQE 17
 | | | | |
 Db 2 ISFDLLSDE 10

RESULT 125
 C49164
 chromogranin-B - rat (fragment)
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-May-1997
 C;Accession: C49164
 R;Nielsen, E.; Welinder, B.S.; Madsen, O.D.
 Endocrinology 129, 3147-3156, 1991
 A;Title: Chromogranin-B, a putative precursor of eight novel rat glucagonoma peptides th
 A;Reference number: A49164; MUID:92063871; PMID:1954895
 A;Accession: C49164
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-20 <NIE>
 A;Note: sequence extracted from NCBI backbone (NCBIP:66369)
 C;Superfamily: chromogranin B precursor

Query Match 22.7%; Score 22; DB 2; Length 20;
 Best Local Similarity 66.7%; Pred. No. 2.9e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EPNHLN 6
 | | | | |
 Db 10 EKNYLN 15

RESULT 126

PT0070

hypothetical protein (gtfc 3' region) - Streptococcus mutans (fragment)
C:Species: Streptococcus mutans
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: PT0070
R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A:Title: Sequence analysis of the gtfc gene from Streptococcus mutans GS-5.
A:Reference number: JT0345; MUID:89137980; PMID:2976010
A:Accession: PT0070
A:Molecule type: DNA
A:Residues: 1-22 <UED>
A:Cross-references: UNIPROT:Q54449; UNIPROT:Q8VUH2
A:Experimental source: strain GS-5

Query Match 22.7%; Score 22; DB 2; Length 22;
Best Local Similarity 29.4%; Pred. No. 3.2e+03;
Matches 5; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQE 17
| | | | : : |
DB 4 EINHLEKVFTRFSKEE 20

RESULT 127

S58242

pyrroloquinoline quinone precursor pqqA - Pseudomonas fluorescens
N:Alternate names: pyrroloquinoline quinone biosynthesis A
C:Species: Pseudomonas fluorescens
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C:Accession: S58242
R:Schneider, U.; Keel, C.; Defago, G.; Haas, D.
submitted to the EMBL Data Library, May 1995
A:Description: Tn5-directed cloning of pqq genes from Pseudomonas fluorescens CHA0: the
A:Reference number: S58239

A:Accession: S58242

A:Molecule type: DNA

A:Residues: 'MRQHSHPPQRSNF', 1-24 <SCH>

A:Cross-references: UNIPROT:P55171; EMBL:X87299; NID:G929799; PIDN:CAA60731.1; PID:G9298

A>Note: in Genbank entry PFPQABCF, release 116.0, the indicated alternative sequence (F

C:Genetics:

A:Gene: pqqA

C:Superfamily: pyrroloquinoline quinone precursor pqqA

C:Keywords: quinoprotein

F16.20/Product: pyrroloquinoline quinone #status predicted <MAT>

F16-20/Cross-link: pyrroloquinoline quinone (Glu, Tyr) #status predicted

Query Match 22.7%; Score 22; DB 1; Length 24;
Best Local Similarity 23.1%; Pred. No. 3.5e+03;
Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPFKI 13
: | : : | : :
DB 5 KPAYTDLRIGFEV 17

RESULT 128

A32203

dihydrofolate reductase (SC 1.5.1.3) - Yeast (Candida albicans) (fragment)
C:Species: Candida albicans
C:Date: 20-Jul-1989 #sequence_revision 20-Jul-1989 #text_change 09-Jul-2004
C:Accession: A32203
R:Baccanari, D.P.; Tansik, R.L.; Joyner, S.S.; Fling, M.E.; Smith, P.L.; Freisheim, J.H.
J. Biol. Chem. 264, 1100-1107, 1989
A:Title: Characterization of Candida albicans dihydrofolate reductase.
A:Reference number: A32203; MUID:89093095; PMID:2642898
A:Accession: A32203

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-25 <BAC>

A:Cross-references: UNIPROT:P22906

C:Keywords: NADP; oxidoreductase

Query Match

22.7%; Score 22; DB 2; Length 25;
Best Local Similarity 26.7%; Pred. No. 3.7e+03;
Matches 4; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSQEPA 19
| : | : | : |
DB 2 LKPNVALLVAALKPA 16

RESULT 129

I51597

albumin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C:Accession: I51597
R:Sweeney, G.; Brooks, A.; Day, P.; Old, R.
Nucleic Acids Res. 15, 5889, 1987
A:Title: DNA sequence of the first exon and 5' flanking region of the 68-K serum albumin
A:Reference number: I51597; MUID:87289055; PMID:3615206
A:Accession: I51597
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-26 <SWE>
A:Cross-references: UNIPROT:P87466; EMBL:Y00381; NID:G64493; PIDN:CAA68453.1; PID:G6449
C:Superfamily: serum albumin; serum albumin repeat homology

Query Match

22.7%; Score 22; DB 2; Length 26;
Best Local Similarity 50.0%; Pred. No. 3.9e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 LNSKIAPK 12
: | : | : | : |
DB 16 IESRIIFK 23

RESULT 130

S55462

mer5 protein homolog - human (fragment)

N:Alternate names: humer

C:Species: Homo sapiens (man)

C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004

C:Accession: S55462

R:Oberbacher, I.

submitted to the EMBL Data Library, March 1995

A:Description: A new member of the highly conserved multigene family of thiol-specific
As by their 3' untranslated regions.

A:Reference number: S55462

A:Accession: S55462

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-30 <OBE>

A:Cross-references: UNIPROT:Q14579; EMBL:X85129; NID:G854125; PIDN:CAA59443.1; PID:G854

C:Genetics:

A:Gene: mer5

C:Superfamily: alkyl hydroperoxide reductase C22 protein; alkyl hydroperoxidase c22 pro

Query Match

22.7%; Score 22; DB 2; Length 30;
Best Local Similarity 37.5%; Pred. No. 4.5e+03;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQ 16
: | : | : | : |
DB 15 KPSPAASKEYFKVQ 30

RESULT 131

A05315

pancreatic ribonuclease (BC 3.1.27.5) K2 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004

C:Accession: A05315

R:Niwata, Y.; Ohgi, K.; Sanda, A.; Takizawa, Y.; Irie, M.

J. Biochem. 97, 923-934, 1985

A;Title: Purification and properties of bovine kidney ribonucleases.

A;Reference number: A91995; MUID:85261183; PMID:3926759

A;Accession: A05315

A;Molecule type: protein

A;Residues: 1-30 <NI>

A;Cross-references: UNIPROT:P08904

A;Experimental source: kidney

C;Superfamily: pancreatic ribonuclease

C;Keywords: hydrolase; kidney; nucleic acid degradation

Query Match 22.7%; Score 22; DB 2; Length 30;
Best Local Similarity 44.4%; Pred. No. 4.5e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EPNHLNSKI 9

Db 12 EIQHIQSR I 20

RESULT 132

A47607

immunogenic protein MPB64 - Mycobacterium bovis (strain BCG) (fragment)

C;Species: Mycobacterium bovis

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999

C;Accession: A47607

R;Harboe, M.; Nagai, S.; Patarroyo, M.E.; Torres, M.L.; Ramirez, C.; Cruz, N.

Infect. Immun. 52, 293-302, 1986

A;Title: Properties of proteins MPB64, MPB70, and MPB80 of Mycobacterium bovis BCG.

A;Reference number: A47607; MUID:86166829; PMID:3514457

A;Accession: A47607

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-30 <HAR>

Query Match 22.7%; Score 22; DB 2; Length 30;
Best Local Similarity 40.0%; Pred. No. 4.5e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 AKFIVSQEPA 19

Db 17 AYQIQMSDPA 26

RESULT 133

F82209

hypothetical protein VC1352 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: F82209

R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I.

l., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: F82209

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-30 <HEI>

A;Cross-references: UNIPROT:Q9KSA7; GB:AE004215; GB:AE003852; NID:g9655842; PIDN:AAF9451

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC1352

A;Map position: 1

Query Match 22.7%; Score 22; DB 2; Length 30;
Best Local Similarity 66.7%; Pred. No. 4.5e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 KIAFKI 13

Db 5 KLGFKI 10

RESULT 134

PH0248

T-cell receptor Vb CDR3, HAM1TCR Vb7a CDR 3a.sbt - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 09-May-1997

C;Accession: PH0248

R;Hara, H.; Morita, M.; Iwaki, T.; Hatae, T.; Itoyama, Y.; Kitamoto, T.; Akizuki, S.; G.

submitted to JIPID, June 1994

A;Description: Detection of HTLV-I proviral DNA and analysis of T cell receptor Vb CDR3

A;Reference number: PH0227

A;Accession: PH0248

A;Molecule type: mRNA

A;Residues: 1-31 <HAR>

A;Experimental source: spinal cord

C;Genetics:

A;Map position: 7

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: receptor

Query Match 22.7%; Score 22; DB 2; Length 31;
Best Local Similarity 80.0%; Pred. No. 4.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 SQEPA 19

Db 5 SQDPA 9

RESULT 135

B82401

hypothetical protein VCA0918 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: B82401

R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I.

l., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: B82401

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-31 <HBI>

A;Cross-references: UNIPROT:Q9KL31; GB:AE004420; GB:AE003853; NID:g9658361; PIDN:AAF9687

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VCA0918

A;Map position: 2

Query Match 22.7%; Score 22; DB 2; Length 31;
Best Local Similarity 60.0%; Pred. No. 4.7e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLSN 6

Db 4 PDHFN 8

RESULT 136

S23476

hypothetical protein 1 (cpg3 3' region) - Fischerella sp. (fragment)

C;Species: Fischerella sp.

C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 15-Oct-1999

C;Accession: S23476; S16061

R;Glauser, M.; Stirewalt, V.L.; Bryant, D.A.; Sidler, W.; Zuber, H.

Eur. J. Biochem. 205, 927-937, 1992

A;Title: Structure of the genes encoding the rod-core linker polypeptides of Mastigoclad

A;Reference number: S23472; MUID:92249337; PMID:1577010

A;Accession: S23476

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-32 <GLA>
 A;Cross-references: EMBL:X59763; NID:g44397; PIDN:CAA42436.1; PID:g44402
 A;Note: the source is designated as Mastigocladus laminosus

Query Match 22.7%; Score 22; DB 2; Length 32;
 Best Local Similarity 25.0%; Pred. No. 4.9e+03;
 Matches 3; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 3 NHLNSKIAPKIV 14
 : : : : :
 Db 16 NYINKLYLVLI 27

RESULT 137

Ti4569
 hypothetical protein 32 - beet chloroplast
 C;Species: chloroplast Beta vulgaris (beet)
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C;Accession: T14569
 R;Kan, Z.; Michaelis, G.
 Theor. Appl. Genet. 91, 836-840, 1995
 A;Title: Mapping of a chloroplast RFLP marker associated with the CMS cytoplasm of sugar
 A;Reference number: Z18144
 A;Accession: T14569
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: DNA
 A;Residues: 1-32 <RAN>
 A;Cross-references: UNIPROT:Q31735; EMBL:X87636; NID:g860887; PID:g860890
 C;Genetics:
 A;Genome: chloroplast
 C;Keywords: chloroplast

Query Match 22.7%; Score 22; DB 2; Length 32;
 Best Local Similarity 40.0%; Pred. No. 4.9e+03;
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HNSKIAPKI 13
 : : : : :
 Db 18 NINSKNKPNL 27

RESULT 138

GNGPB
 big gastrin [validated] - guinea pig
 N;Contains: gastrin
 C;Species: Cavia porcellus (guinea pig)
 C;Date: 30-Sep-1987 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
 C;Accession: A26089
 R;Bonato, C.; Eng, J.; Pan, Y.C.E.; Miedel, M.; Hulmes, J.D.; Yalow, R.S.
 Life Sci. 39, 959-964, 1986
 A;Title: Guinea pig 33-amino acid gastrin.
 A;Reference number: A26089; MUID:86309993; PMID:3747718
 A;Accession: A26089
 A;Molecule type: protein
 A;Residues: 1-33 <BON>
 A;Cross-references: UNIPROT:P06885
 C;Comment: Big gastrin constitutes only about 5% of antral gastrin.

C;Superfamily: gastrin
 C;Keywords: amidated carboxyl end; hormone; pyroglutamic acid
 F1-33/Product: big gastrin #status experimental <BGN>
 F18-33/Product: gastrin #status experimental <SGN>
 F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F133/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 22.7%; Score 22; DB 1; Length 33;
 Best Local Similarity 36.4%; Pred. No. 5e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PNHLNSKIAPK 12
 : : : : :
 Db 7 PAHLRTDLSKK 17

RESULT 139

B29541
 big gastrin - Chinchilla brevicaudata
 C;Species: Chinchilla brevicaudata, Chinchilla lanigera brevicaudata
 C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
 C;Accession: B29541
 R;Shinomura, Y.; Eng, J.; Yalow, R.S.
 Biochem. Biophys. Res. Commun. 143, 7-14, 1987
 A;Title: Chinchilla "big" and "little" gastrins
 A;Reference number: A90130; MUID:87156784; PMID:3827930
 A;Accession: B29541
 A;Molecule type: protein
 A;Residues: 1-33 <SHI>
 A;Cross-references: UNIPROT:P10034
 C;Superfamily: gastrin

Query Match 22.7%; Score 22; DB 2; Length 33;
 Best Local Similarity 36.4%; Pred. No. 5e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PNHLNSKIAPK 12
 : : : : :
 Db 7 PPHLGTDLSSK 17

RESULT 140

161695
 myosin - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 09-Jul-2004
 C;Accession: I61695
 R;Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.
 Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994
 A;Title: Identification and overlapping expression of multiple unconventional myosin genes
 A;Reference number: A55758; MUID:94294418; PMID:8022818
 A;Accession: I61695
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: mRNA
 A;Residues: 1-33 <RES>
 A;Cross-references: UNIPROT:Q14783; GB:L29144; NID:g457252; PIDN:AAA20907.1; PID:g53113
 C;Superfamily: myosin MYO2; myosin motor domain homology

Query Match 22.7%; Score 22; DB 2; Length 33;
 Best Local Similarity 23.1%; Pred. No. 5e+03;
 Matches 3; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 6 NSKIAPKIVSQEP 18
 : : : : :
 Db 19 NAHVEDKVLASNP 31

RESULT 141

146596
 myosin - pig (fragment)
 C;Species: Sus scrofa domestica (domestic pig)
 C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
 C;Accession: I46596
 R;Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.
 Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994
 A;Title: Identification and overlapping expression of multiple unconventional myosin genes
 A;Reference number: A55758; MUID:94294418; PMID:8022818
 A;Accession: I46596
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: mRNA
 A;Residues: 1-33 <BEM>
 A;Cross-references: UNIPROT:Q29062; GB:L29131; NID:g457341; PIDN:AAA20916.1; PID:g53114
 C;Superfamily: myosin MYO2; myosin motor domain homology

Query Match 22.7%; Score 22; DB 2; Length 33;
 Best Local Similarity 23.1%; Pred. No. 5e+03;
 Matches 3; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 6 NSKIAPKIVSQEP 18

Db 19 NAHVEDKVLASNP 31

RESULT 142

T08018

ycf12 protein - Chlamydomonas reinhardtii chloroplast

C;Species: Chloroplast Chlamydomonas reinhardtii

C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004

C;Accession: T08018

R;Khrebtukova, I.; Spreitzer, R.J.

submitted to the EMBL Data Library, November 1995

A;Reference number: Z16296

A;Accession: T08018

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-33 <KHR>

A;Cross-references: UNIPROT:P03070; EMBL:U40346; NID:g1101912; PIDN:AAA91171.1; PID:g1101912

C;Genetics:

A;Gene: ycf12

A;Genome: chloroplast

C;Keywords: chloroplast

Query Match 22.7%; Score 22; DB 2; Length 33;

Best Local Similarity 36.4%; Pred. No. 5e+03;

Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAPFKIVS 15

Db 1 MNIELALTIVS 11

RESULT 143

S35572

zona pellucida protein alpha chain - pig

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 20-May-1994 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004

C;Accession: S35572

R;Toepfer-Petersen, E.; Mann, K.; Calvete, J. J.

Biol. Chem. Hoppe-Seyler 374, 411-417, 1993

A;Title: Identification of porcine oocyte 55 kDa alpha and beta proteins within the zona

a pellucida proteins in different mammalian species.

A;Reference number: S35572; MUID:94030657; PMID:8216892

A;Accession: S35572

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-36 <TOE>

A;Cross-references: UNIPROT:Q7M398

C;Superfamily: sperm-binding glycoprotein 2P3-alpha; trefoil homology; ZP domain homolog

Query Match 22.7%; Score 22; DB 2; Length 36;

Best Local Similarity 30.8%; Pred. No. 5.5e+03;

Matches 4; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPNHLSNKIAFKI 13

Db 21 EPIYVEVSIRFSV 33

RESULT 144

D82457

hypothetical protein VCA0452 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: D82457

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82457

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-36 <HEI>

A;Cross-references: UNIPROT:Q9XWC0; GB:AE004378; GB:AE003853; NID:g9657851; PIDN:AAF963

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VCA0452

A;Map position: 2

Query Match 22.7%; Score 22; DB 2; Length 36;

Best Local Similarity 38.9%; Pred. No. 5.5e+03;

Matches 7; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

QY 5 LNSKI-----AFKIVSQ 16

Db 2 LNEKVSVMKMLTFKNVSR 19

RESULT 145

S33435

hypothetical protein L - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: S33435

R;Fischer, R.J.; Helms, J.; Duerre, P.

submitted to the EMBL Data Library, May 1993

A;Description: Cloning, sequencing and mRNA analysis of the sol operon from Clostridium

A;Reference number: S33432

A;Accession: S33435

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-36 <FIC>

A;Cross-references: UNIPROT:Q45814; EMBL:X72831; NID:g298080; PIDN:CAA51343.1; PID:g298080

Query Match 22.7%; Score 22; DB 2; Length 36;

Best Local Similarity 83.3%; Pred. No. 5.5e+03;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 SKIAFK 12

Db 16 SKSAFK 21

RESULT 146

HSPV5

histone H5 - pigeon (fragment)

C;Species: Columba livia (domestic pigeon)

C;Date: 31-May-1980 #sequence_revision 31-May-1980 #text_change 09-Jul-2004

C;Accession: A02590

R;Yaguchi, M.; Roy, C.; Dove, M.; Seligy, V.

Biochem. Biophys. Res. Commun. 76, 100-106, 1977

A;Title: Amino acid sequence homologies between H1 and H5 histones.

A;Reference number: A02590; MUID:77201524; PMID:559492

A;Accession: A02590

A;Molecule type: protein

A;Residues: 1-38 <YAG>

A;Cross-references: UNIPROT:P02260

C;Superfamily: histone H1

C;Keywords: chromosomal protein; DNA binding; erythrocyte

F;1-23/Domain: amino-terminal <NH2>

F;24-38/Domain: globular (fragment) <GLB>

Query Match 22.7%; Score 22; DB 1; Length 38;

Best Local Similarity 62.5%; Pred. No. 5.9e+03;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 KIVSQEPA 19

Db 19 KIVSQEPA 26

RESULT 147

S70800

Ivi protein III - Vibrio cholerae (fragment)

A;Cross-references: GB:AL513382; PIDN:CAD08037.1; PID:gl6505019; GSPDB:GN00117
C;Genetics:
A;Gene: STY4216

Query Match 22.7%; Score 22; DB 2; Length 38;
Best Local Similarity 37.5%; Pred. No. 5.9e+03;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0

QY 11 FKIVSQEP 18
:|:|:
Db 21 YKVAVKVP 28

RESULT 150
A44862
microtubule associated protein (MAP) homolog - Trypanosoma cruzi (fragment)
C;Species: Trypanosoma cruzi
C;Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A44862
E;Kerner, N.; Liegeard, P.; Levin, M.J.; Hontebeyrie-Joskowicz, M.
Exp. Parasitol. 73, 451-459, 1991
A;Title: Trypanosoma cruzi: antibodies to a MAP-like protein in chronic Chagas
A;Reference number: A44862; MUID:92070444; PMID:1959572
A;Accession: A44862
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-38 <R>
A;Cross-references: UNIPROT:Q9U9R2; GB:S68286; MID:g2339898; PID:g2339899
A;Experimental source: Epimastigotes, Tulahuen 2 strain
A;Note: sequence extracted from NCBI backbone (NCBIN:68286, NCBIIP:68287)

Query Match 22.7%; Score 22; DB 2; Length 38;
Best Local Similarity 42.9%; Pred. No. 5.9e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0

QY 1 EPNHLS 7
:|:|:
Db 15 DPDHFRS 21

RESULT 151
S10315
photosystem II protein psbI - Synechococcus sp.
C;Species: Synechococcus sp.
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-Aug-1999
C;Accession: S10315
E;Chen, J.C.; Meng, B.Y.; Fukuta, M.; Sugiura, M.
Nucleic Acids Res. 18, 4017, 1990
A;Title: Nucleotide sequence of the psbI gene of the cyanobacterium, Anacystis
A;Reference number: S10315; MUID:90326553; PMID:2115674
A;Accession: S10315
A;Molecule type: DNA
A;Residues: 1-39 <CHE>
A;Cross-references: EMBL:X52750; NID:g38910; PIDN:CAA36961.1; PID:g38911
C;Genetics:
A;Gene: psbI
C;Superfamily: photosystem II protein psbI
C;Keywords: membrane-associated complex; photosynthesis; photosystem II; thylakoid

Query Match 22.7%; Score 22; DB 2; Length 39;
Best Local Similarity 44.4%; Pred. No. 6.1e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0

QY 11 FKIVSQEPA 19
|:|:|:
Db 21 FGFLSDPA 29

RESULT 152
S67938
hypothetical protein 3 - Streptococcus pneumoniae
C;Species: Streptococcus pneumoniae
C;Date: 05-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997

C;Accession: S67938
R;Saluja, S.K.; Weiser, J.N.
Mol. Microbiol. 16, 215-227, 1995
A;Title: The genetic basis of colony opacity in *Streptococcus pneumoniae*: evidence for a
A;Reference number: S67936; MUID:96015435; PMID:7565084
A;Accession: S67938
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-39 <SAL>
A;Cross-references: EMBL:U12567

Query Match 22.7%; Score 22; DB 2; Length 39;
Best Local Similarity 38.5%; Pred. No. 6.1e+03;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NMLNSKIAFKIVS 15
Db 14 NEVNLLIKRIIS 26

RESULT 153
G82287
hypothetical protein VC0733 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C;Species: *Vibrio cholerae*
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82287
R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Makalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82287
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-39 <HEI>
A;Cross-references: UNIPROT:Q9KTYZ7; GB:AE004159; GB:AE003852; NID:G9655167; PIDN:AAF9389
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0733
A;Map position: 1

Query Match 22.7%; Score 22; DB 2; Length 39;
Best Local Similarity 26.7%; Pred. No. 6.1e+03;
Matches 4; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NMLNSKIAFKIVSOE 17
Db 17 NLITQKLTALKQ 31

RESULT 154
D61320
plastocyanin - *Consolida ajacis* (fragment)
C;Species: *Consolida ajacis*
C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
C;Accession: D61320
R;Grund, C.; Gilroy, J.; Gleaves, T.; Jensen, U.; Boulter, D.
Phytochemistry 20, 1559-1565, 1981
A;Title: Systematic relationships of the Ranunculaceae based on amino acid sequence data
A;Reference number: A61320
A;Accession: D61320
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-40 <GRU>
A;Cross-references: UNIPROT:Q7M256
C;Superfamily: plastocyanin
C;Keywords: electron transfer; metalloprotein

Query Match 22.7%; Score 22; DB 2; Length 40;
Best Local Similarity 40.0%; Pred. No. 6.2e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

Qy 2 PNHL-----NSKIAFK 12
Db 16 PNNFTVSAGEKIVFK 30

RESULT 155
AH2446
hypothetical protein asl5128 [imported] - *Nostoc* sp. (strain PCC 7120)
C;Species: *Nostoc* sp. PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2446
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2446
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-40 <KUR>
A;Cross-references: UNIPROT:Q8YM13; GB:BA000019; PIDN:BAB76827.1; PID:gl7134266; GSPDB:
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asl5128

Query Match 22.7%; Score 22; DB 2; Length 40;
Best Local Similarity 25.0%; Pred. No. 6.2e+03;
Matches 3; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HLNSKIAFKIVS 15
Db 7 NINWEVIFQUTS 18

RESULT 156
S71301
ICL5 protein - *Paramecium tetraurelia* (fragment)
C;Species: *Paramecium tetraurelia*
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-Dec-1999
C;Accession: S71301
R;Madeddu, L.; Klotz, C.; le Caer, J.P.; Beisson, J.
Eur. J. Biochem. 238, 121-128, 1996
A;Title: Characterization of centrin genes in *Paramecium*.
A;Reference number: S71298; MUID:96248429; PMID:8665928
A;Accession: S71301
A;Molecule type: protein
A;Residues: 1-40 <MAD>
A;Experimental source: strain d4-2
C;Genetics:
A;Genetic code: SGCS

Query Match 22.7%; Score 22; DB 2; Length 40;
Best Local Similarity 66.7%; Pred. No. 6.2e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 14 VSQEPA 19
Db 2 VEQPPA 7

RESULT 157
S44136
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - yeast (*Candida parapsilosis*) mit
C;Species: *Candida parapsilosis*
C;Date: 06-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 03-Jun-2002
C;Accession: S44136
R;Nosek, J.
submitted to the EMBL Data Library, November 1993
A;Reference number: S44135
A;Accession: S44136
A;Molecule type: DNA
A;Residues: 1-40 <NOS>

A;Cross-references: EMBL:X75676; NID:g473032; PIDN:CAA53373.1; PID:g996061

C;Genetics:

A;Gene: ND2

A;Genome: mitochondrion

A;Genetic code: SGC3

C;Keywords: mitochondrion; NAD; oxidoreductase

Query Match 22.7%; Score 22; DB 2; Length 40;
Best Local Similarity 44.4%; Pred. No. 6.2e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAFKI 13

Db 6 INKSLAFVI 14

RESULT 158

H82330

hypothetical protein VC0380 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: H82330

R;Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000.

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: H82330

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-40 <HEI>

A;Cross-references: UNIPROT:Q9KUX8; GB:AE004126; GB:AE003852; NID:g9654802; PIDN:AAF9359

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC0380

A;Map position: 1

Query Match 22.7%; Score 22; DB 2; Length 40;
Best Local Similarity 80.0%; Pred. No. 6.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHL 5

Db 12 EANHL 16

RESULT 159

F64009

hypothetical protein HI0557 - Haemophilus influenzae (strain Rd KW20)

C;Species: Haemophilus influenzae

C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C;Accession: F64009

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A;Reference number: A64000; MUID:95350630; PMID:7542800

A;Accession: F64009

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-40 <TIGR>

A;Cross-references: UNIPROT:P44015; GB:U32737; GB:LA42023; NID:g1573536; PIDN:AAC22218.1;

Query Match 22.7%; Score 22; DB 2; Length 40;

Best Local Similarity 53.3%; Pred. No. 6.2e+03;

Matches 8; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 4 HLNSKIA--FKIVSQ 16

Db 11 HAVDKIAQEFVIYSQ 25

RESULT 160

PQ0564

nonstructural protein 3 (clone 4) - hepatitis C virus (fragment)

C;Species: hepatitis C virus

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004

C;Accession: PQ0564

R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno

Biochem. Biophys. Res. Commun. 181, 279-285, 1991

A;Title: Distribution of plural HCV types in Japan.

A;Reference number: PQ0554; MUID:92068204; PMID:1720309

A;Accession: PQ0564

A;Molecule type: mRNA

A;Residues: 1-41 <KAT>

A;Cross-references: UNIPROT:O81251; GB:D10567; GB:D90523; NID:g221537; PIDN:BAA01423.1;

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: nonstructural protein

Query Match 22.7%; Score 22; DB 2; Length 41;
Best Local Similarity 50.0%; Pred. No. 6.4e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 HLNSKI 9

Db 27 HLNDQV 32

RESULT 161

PQ0565

nonstructural protein 3 (clone 13) - hepatitis C virus (fragment)

C;Species: hepatitis C virus

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004

C;Accession: PQ0565

R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno

Biochem. Biophys. Res. Commun. 181, 279-285, 1991

A;Title: Distribution of plural HCV types in Japan.

A;Reference number: PQ0554; MUID:92068204; PMID:1720309

A;Accession: PQ0565

A;Molecule type: mRNA

A;Residues: 1-41 <KAT>

A;Cross-references: UNIPROT:O81249; GB:D10569; GB:D90524; NID:g221535; PIDN:BAA01424.1;

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: nonstructural protein

Query Match 22.7%; Score 22; DB 2; Length 41;
Best Local Similarity 50.0%; Pred. No. 6.4e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 HLNSKI 9

Db 27 HLNDQV 32

RESULT 162

T07263

hypothetical protein 41b - Chlorella vulgaris chloroplast

C;Species: chloroplast Chlorella vulgaris

C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000

C;Accession: T07263

R;Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Naka

Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997

A;Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chl

A;Reference number: Z15985; MUID:97303241; PMID:9159184

A;Accession: T07263

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-41 <WAX>

A;Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57911.1; PID:g2224427

C;Genetics:

A;Genome: chloroplast

C;Keywords: chloroplast

Query Match 22.7%; Score 22; DB 2; Length 41;
 Best Local Similarity 55.6%; Pred. No. 6.4e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAPKI 13
 :||| :|||
 Db 1 MNSKYFFLI 9

RESULT 163
 A30010
 Cytochrome-c oxidase (EC 1.9.3.1) chain II homolog - Leishmania tarentolae mitochondrion
 C;Species: mitochondrion Leishmania tarentolae
 C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 18-Aug-2003
 C;Accession: A30010
 R;de la Cruz, V.F.; Neckelmann, N.; Simpson, L.
 J. Biol. Chem. 259, 15136-15147, 1984
 A;Title: Sequences of six genes and several open reading frames in the kinetoplast maxic
 A;Reference number: A22848; MUID:8507995; PMID:6096360
 A;Accession: A30010
 A;Molecule type: DNA
 A;Residues: 1-42
 A;Cross-references: GB:M10126
 C;Genetics:
 A;Genome: mitochondrion
 A;Genetic code: SGC6
 C;Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase c
 C;Keywords: copper; electron transfer; membrane-associated complex; mitochondrial inner

Query Match 22.7%; Score 22; DB 2; Length 42;
 Best Local Similarity 36.4%; Pred. No. 6.6e+03;
 Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPNHLSKIAP 11
 ||| :|||
 Db 3 EPGRCNEIVLF 13

RESULT 164
 S01125
 photosystem II phosphoprotein psbH - Chlamydomonas reinhardtii (fragment)
 N;Alternate names: photosystem II 8K phosphoprotein
 C;Species: Chlamydomonas reinhardtii
 C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
 C;Accession: S01125
 R;Dedner, N.; Meyer, H.E.; Ashton, C.; Wildner, G.F.
 FEBS Lett. 236, 77-82, 1988
 A;Title: N-terminal sequence analysis of the 8 kDa protein in Chlamydomonas reinhardtii.
 A;Reference number: S01125
 A;Accession: S01125
 A;Molecule type: protein
 A;Residues: 1-42 <DED>
 A;Cross-references: UNIPROT:P22666
 C;Genetics:
 A;Gene: psbH
 C;Superfamily: photosystem II phosphoprotein psbH
 C;Keywords: chloroplast; phosphoprotein; photosynthesis; photosystem II; transmembrane P
 F;2/Binding site: phosphate (Thr) (covalent) #status experimental

Query Match 22.7%; Score 22; DB 2; Length 42;
 Best Local Similarity 33.3%; Pred. No. 6.6e+03;
 Matches 6; Conservative 3; Mismatches 3; Indels 6; Gaps 1;

QY 1 EPNHLSKIAPKIVSOEP 18
 :||: :|||
 Db 8 KPSKVNDSF-----QEP 19

RESULT 165
 PQ0182
 naringenin-chalcone synthase (EC 2.3.1.74) 2 - garden pea (fragment)
 N;Alternate names: chalcone synthase
 C;Species: Pisum sativum (garden pea)

C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 09-Jul-2004
 C;Accession: PQ0182
 R;Harker, C.L.; Ellis, T.H.N.; Coen, E.S.
 Plant Cell 2, 185-194, 1990
 A;Title: Identification and genetic regulation of the chalcone synthase multigene famil
 A;Reference number: PQ0181; MUID:93005652; PMID:2152111
 C;Accession: PQ0182
 A;Molecule type: mRNA
 A;Residues: 1-44 <HAR>
 A;Cross-references: UNIPROT:Q9S9B9
 A;Note: the gene encoding this protein is expressed in roots
 C;Comment: This enzyme catalyzes the first step in phenylpropanoid biosynthesis.
 C;Genetics:
 A;Gene: CHS2
 C;Superfamily: Type III polyketide synthase
 C;Keywords: acyltransferase; coenzyme A

Query Match 22.7%; Score 22; DB 2; Length 44;
 Best Local Similarity 50.0%; Pred. No. 6.9e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 AFKIVSQE 17
 :||| :|||
 Db 35 SEKITNSE 42

RESULT 166
 T26893
 hypothetical protein Y44A6D.1 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T26893
 R;Ainscough, R.
 submitted to the EMBL Data Library, June 1998
 A;Reference number: Z20282
 A;Accession: T26893
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-44 <WIL>
 A;Cross-references: UNIPROT:Q9XXE3; EMBL:AL023842; PIDN:CAA19514.1; GSPDB:GN00023; CESP:
 A;Experimental source: clone Y44A6D
 C;Genetics:
 A;Gene: CESP:Y44A6D.1
 A;Map position: 5
 A;Introns: 30/1

Query Match 22.7%; Score 22; DB 2; Length 44;
 Best Local Similarity 37.5%; Pred. No. 6.9e+03;
 Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 HLNSKIAP 11
 ||| :|||
 Db 8 HLHTALAY 15

RESULT 167
 H70249
 hypothetical protein BBJ39 - Lyme disease spirochete plasmid J/1p38
 C;Species: Borrelia burgdorferi (Lyme disease spirochete)
 C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 C;Accession: H70249
 R;Fraser, C.W.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kurlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A;Authors: Smith, H.O.; Venter, J.C.
 A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
 A;Reference number: A70100; MUID:98065943; PMID:9403685
 A;Accession: H70249
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-45 <KLE>
 A;Cross-references: UNIPROT:O50794; GB:AE000787; NID:g2690175; PIDN:AAC66113.1; PID:g269

A;Experimental source: strain B31
C;Genetics:
A;Genome: plasmid

Query Match 22.7%; Score 22; DB 2; Length 45;
Best Local Similarity 37.5%; Pred. No. 7.1e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy 9 IAFKIVSQ 16
Db 12 ITFSVISE 19

RESULT 168

H83936
hypothetical protein BH2296 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H83936
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H83936
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-45 <STO>
A;Cross-references: UNIPROT:Q9KXJ1; GB:AP001515; GB:BA000004; NID:gi10174886; PIDN:BA060
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2296

Query Match 22.7%; Score 22; DB 2; Length 45;
Best Local Similarity 33.3%; Pred. No. 7.1e+03;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 1 EPNHLSKI 9
Db 35 EQRHISKXL 43

RESULT 169

D82363

hypothetical protein VC0102 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: D82363
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: D82363
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-46 <HE1>
A;Cross-references: UNIPROT:Q9KVN9; GB:AE004101; GB:AE003852; NID:gi9654497; PIDN:AAF9328
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0102
A;Map position: 1

Query Match 22.7%; Score 22; DB 2; Length 46;
Best Local Similarity 66.7%; Pred. No. 7.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 3 NHLSK 8
Db 14 NHLQK 19

RESULT 170

G83936

hypothetical protein BH2295 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: G83936
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G83936
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-46 <STO>
A;Cross-references: UNIPROT:Q9KXJ2; GB:AP001515; GB:BA000004; NID:gi10174886; PIDN:BA060
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2295

Query Match 22.7%; Score 22; DB 2; Length 46;
Best Local Similarity 42.9%; Pred. No. 7.3e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 2 PNHLSK 8
Db 13 PDELGTQ 19

RESULT 171

E84357

50S ribosomal protein L40E [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: E84357
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jahn
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lasky,
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: E84357
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-47 <STO>
A;Cross-references: UNIPROT:Q9HNJ5; GB:AE004437; NID:gi0581494; PIDN:AG20225.1; GSPD:
C;Genetics:
A;Gene: rpl40e

Query Match 22.7%; Score 22; DB 2; Length 47;
Best Local Similarity 22.2%; Pred. No. 7.4e+03;
Matches 4; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Qy 1 EPNHLSKIAPKIVSQEP 18
Db 6 EDRLLNKQVCMRCNARNP 23

RESULT 172

A82385

hypothetical protein VC0104 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82385
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers,
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82385
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-47 <HE1>

A;Cross-references: UNIPROT:Q9KKQ8; GB:AE004430; GB:AE003853; NID:g9658484; PIDN:AAF9693
A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor
C;Genetics:
A;Gene: VCA1044
A;Map position: 2

Query Match 22.7%; Score 22; DB 2; Length 47;
Best Local Similarity 50.0%; Pred. No. 7.4e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HLNSKIAFKI 13
::|||
DB 12 YINLUKIIDKI 21

RESULT 173
H83869
hypothetical protein BH1760 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H83869
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H83869
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-47 <STO>
A;Cross-references: UNIPROT:Q9KC14; GB:AF001513; GB:BA000004; NID:g10174345; PIDN:BA054
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1760

Query Match 22.7%; Score 22; DB 2; Length 47;
Best Local Similarity 44.4%; Pred. No. 7.4e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 AKFIVSQEP 18
|||:
DB 33 ALKALKKEP 41

RESULT 174
PWCK8P
H+-transporting two-sector ATPase (EC 3.6.3.14) protein 8 - yeast (Candida parapsilosis)
C;Species: mitochondrion Candida parapsilosis
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S10465
R;Guelin, E.; Velours, J.; Guerin, M.
Nucleic Acids Res. 18, 4267, 1990
A;Title: Cloning and sequencing of a fragment of the linear mitochondrial DNA of the yeast
A;Reference number: S10465; MUID:90332440; PMID:2143015
A;Accession: S10465
A;Molecule type: DNA
A;Residues: 1-48 <GUE>
A;Cross-references: UNIPROT:P17345; EMBL:X52115; NID:g12890; PIDN:CAA36361.1; PID:g82933
C;Genetics:
A;Gene: aap1
A;Genome: mitochondrion
A;Genetic code: SGC3
C;Superfamily: yeast H+-transporting ATP synthase protein 8
C;Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion; oxid

Query Match 22.7%; Score 22; DB 1; Length 48;
Best Local Similarity 46.2%; Pred. No. 7.6e+03;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 PNHLNSKIAFKIV 14
|||:
DB 33 PNILRLLIARNII 45

RESULT 175

B86779
hypothetical protein ymgB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86779
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrh
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86779
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-48 <STO>
A;Cross-references: UNIPROT:Q9CG76; GB:AE005176; PID:g12724205; PIDN:AAK05332.1; GSPDB:
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ymgB

Query Match 22.7%; Score 22; DB 2; Length 48;
Best Local Similarity 66.7%; Pred. No. 7.6e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NHLNSK 8
|||
DB 11 NHANIK 16

RESULT 176

S61469
p83/100 protein - Borrelia afzelii (strain Ple and others) (fragment)
C;Species: Borrelia afzelii
A;Variety: strain Ple and others
C;Date: 19-Mar-1997 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: S61469; S61470; S61471
R;Roessler, D.; Eifert, H.; Jauris-Heipke, S.; Lehnert, G.; Preac-Mursic, V.; Teepe, J.
Med. Microbiol. Immunol. 184, 23-32, 1995
A;Title: Molecular and immunological characterization of the p83/100 protein of various
A;Reference number: S61461; MUID:96149106; PMID:8538575
A;Accession: S61469
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-48 <ROE>
A;Cross-references: UNIPROT:Q57397; EMBL:X81529; NID:g928973; PIDN:CAA57248.1; PID:g928
A;Experimental source: strain Ple; strain PKJ7; strain PGau
C;Keywords: surface antigen

Query Match 22.7%; Score 22; DB 2; Length 48;
Best Local Similarity 62.5%; Pred. No. 7.6e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 LNSKIAFK 12
|||:
DB 22 LNSKASKK 29

RESULT 177

S61472
p83/100 protein - Borrelia afzelii (strain PWudI) (fragment)
C;Species: Borrelia afzelii
A;Variety: strain PWudI
C;Date: 19-Mar-1997 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C;Accession: S61472; S72307
R;Roessler, D.; Eifert, H.; Jauris-Heipke, S.; Lehnert, G.; Preac-Mursic, V.; Teepe, J.
Med. Microbiol. Immunol. 184, 23-32, 1995
A;Title: Molecular and immunological characterization of the p83/100 protein of various
A;Reference number: S61461; MUID:96149106; PMID:8538575
A;Accession: S61472
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-48 <ROE>
A;Cross-references: UNIPROT:Q44673; EMBL:X81535
A;Experimental source: strain PWudI

```
R;Roessler, D.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72307
A;Accession: S72307
A;Molecule type: DNA
A;Residues: 1-7,'F',8-48 <RCW>
A;Cross-references: EMBL:X81535; NID:g928977; PIDN:CAA57254.1; PID:g928978
A;Experimental source: strain Fwudi
C;Keywords: surface antigen

Query Match      22.7%; Score 22; DB 2; Length 48;
Best Local Similarity 62.5%; Pred. No. 7.6e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LNSKIAPK 12
Db 22 LNSKASSK 29

RESULT 178
B48396
ribosomal protein L33 - Bacillus stearothermophilus
C;Species: Bacillus stearothermophilus
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B48396
R;Kruft, V.; Kapp, U.; Wittmann-Liebold, B.
Biochimie 73, 855-860, 1991
A;Title: Characterization and primary structure of proteins L28, L33 and L34 from Bacillus stearothermophilus
A;Reference number: A48396; MUID:92075758; PMID:1742360
A;Accession: B48396
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-49 <RU>
A;Cross-references: UNIPROT:P23375
A;Note: sequence extracted from NCBI backbone (NCBIP:69663)
C;Superfamily: Escherichia coli ribosomal protein L33

Query Match      22.7%; Score 22; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 7.8e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLNSK 8
Db 14 ERNYITSK 21

RESULT 179
A64019
hypothetical protein HI1039 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 10-Oct-1997
C;Accession: A64019
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: A64019
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-50 <TIGR>
A;Cross-references: GB:U32784; GB:U42023; NID:g1574064; PID:g1574074; TIGR:HI1039
C;Genetics:
A;Start codon: GTG

Query Match      22.7%; Score 22; DB 2; Length 50;
Best Local Similarity 50.0%; Pred. No. 8e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 PNHLNSKIAPK 13
Db 14 ERNYITSK 21
```

```
R;Roessler, D.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72307
A;Accession: S72307
A;Molecule type: DNA
A;Residues: 1-7,'F',8-48 <RCW>
A;Cross-references: EMBL:X81535; NID:g928977; PIDN:CAA57254.1; PID:g928978
A;Experimental source: strain Fwudi
C;Keywords: surface antigen

Query Match      22.7%; Score 22; DB 2; Length 48;
Best Local Similarity 62.5%; Pred. No. 7.6e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LNSKIAPK 12
Db 22 LNSKASSK 29

RESULT 178
B48396
ribosomal protein L33 - Bacillus stearothermophilus
C;Species: Bacillus stearothermophilus
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B48396
R;Kruft, V.; Kapp, U.; Wittmann-Liebold, B.
Biochimie 73, 855-860, 1991
A;Title: Characterization and primary structure of proteins L28, L33 and L34 from Bacillus stearothermophilus
A;Reference number: A48396; MUID:92075758; PMID:1742360
A;Accession: B48396
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-49 <RU>
A;Cross-references: UNIPROT:P23375
A;Note: sequence extracted from NCBI backbone (NCBIP:69663)
C;Superfamily: Escherichia coli ribosomal protein L33

Query Match      22.7%; Score 22; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 7.8e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLNSK 8
Db 14 ERNYITSK 21

RESULT 179
A64019
hypothetical protein HI1039 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 10-Oct-1997
C;Accession: A64019
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: A64019
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-50 <TIGR>
A;Cross-references: GB:U32784; GB:U42023; NID:g1574064; PID:g1574074; TIGR:HI1039
C;Genetics:
A;Start codon: GTG

Query Match      22.7%; Score 22; DB 2; Length 50;
Best Local Similarity 50.0%; Pred. No. 8e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 PNHLNSKIAPK 13
Db 14 ERNYITSK 21
```

```
Db 9 PQTLLTKSAVKI 20

RESULT 180
H90596
hypothetical protein MYPV 6800 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: H90596
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Sanson, D.; Galisson, F.; Moszer, I.
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: H90596
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <KOR>
A;Cross-references: UNIPROT:Q98PPO; GB:AL445566; PID:g14090095; PIDN:CAC13853.1; GSPDB:
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPV 6800
A;Genetic code: SGC3

Query Match      22.7%; Score 22; DB 2; Length 50;
Best Local Similarity 28.6%; Pred. No. 8e+03;
Matches 4; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 5 LNSKIAPKIVSQEP 18
Db 26 LKTKVILIIYIQRP 39

RESULT 181
A37968
neural surface protein Bravo - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 15-Aug-1997
C;Accession: A37968; A36345
R;de la Rosa, E.J.; Kayyem, J.F.; Roman, J.M.; Stierhof, Y.D.; Dreyer, W.J.; Schwarz, U.
J. Cell Biol. 112, 1049, 1991
A;Reference number: A37968; MUID:91154309; PMID:1999455
A;Contents: erratum
A;Accession: A37968
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-19 <DEL>
R;de la Rosa, E.J.; Kayyem, J.F.; Roman, J.M.; Stierhof, Y.D.; Dreyer, W.J.; Schwarz, U.
J. Cell Biol. 111, 3087-3096, 1990
A;Title: Topologically restricted appearance in the developing chick retinotectal system
A;Reference number: A36345; MUID:91100421; PMID:2269667
A;Accession: A36345
A;Molecule type: protein
A;Residues: 1-7,9-19 <DEL>

Query Match      22.2%; Score 21.5; DB 2; Length 19;
Best Local Similarity 42.9%; Pred. No. 3.3e+03;
Matches 6; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy 5 LNSKIAPKIVSQEP 18
Db 5 LDSKLIX-EXLSQPP 17

RESULT 182
T48881
leader peptide [imported] - Vibrio sp.
C;Species: Vibrio sp.
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T48881
R;Xu, Y.; Zhang, Y.; Liang, Z.Y.; Van de Castele, M.; Legrain, C.; Glansdorff, N.
Microbiology 144, 1435-1441, 1998
A;Title: Aspartate carbamoyltransferase from a psychrophilic deep-sea bacterium, Vibrio
A;Reference number: Z24845
```

A;Accession: T48881
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Molecule type: DNA
A;Residues: 1-20 <XUY>
A;Cross-references: UNIPROT:P96173; EMBL:Y09786; PI
A;Experimental source: strain 2693

Query Match 22.2%; Score 21.5; DB 2; Length 20;
Best Local Similarity 46.2%; Pred. No. 3.5e+03;
Matches 6; Conservative 3; Mismatches 1; Indels

Qy 2 PNHLNSKIAFKIV 14
 | : | : | : |
Db 6 PSSLS--FKLV 15

RESULT 183

S28940
cyclic nucleotide-binding phosphatase - potato
C:Species: Solanum tuberosum (potato)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S28940
R:Polya, G.M.; Wettenthal, R.E.H.
Biochim. Biophys. Acta 1159, 179-184, 1992
A:Title: Rapid purification and N-terminal
sequencing of a potato tuber cyclic
nucleotide phosphatase
A:Reference number: S28940; MUID:93003376; PMID:11382614
A:Accession: S28940
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-27 <POL>
A:Cross-references: UNIPROT:O988V4

Query Match	22.2%
Query Match	22.2%

Best Local Similarity	60.0%	Pred. No. 4.9e+03	Score 21.5	DB 2	length 27
Matches	6	Conservative	2	Mismatches	1
					Indels

Qy 1 EPNHL-NSKI 9

14 ETNNLQNSKL 23

RESULT 184

S37683
protein IEF SSP 9124 ~ human (fragments)
C:Species: Homo sapiens (man)
C:Date: 09-Dec-1993 #sequence_revision 17-Nov-1995 #text_change 17-Nov-1995
C:Accession: S37683
R:Leffers, H.; Madsen, P.; Rasmussen, H.H.; Honore, B.; Andersen, A.H.; Walbum, E.; Vanden
J. Mol. Biol. 231, 982-998, 1993
A:Title: Molecular cloning and expression of the transformation sensitive epithelial mar
A:Reference number: S34753; MUID:93294871; PMID:8515476
A:Accession: S37683
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-28 <LEF>

Query Match	22.2%;	Score 21.5;	DB 2;	Length 28;
Best Local Similarity	46.2%;	Pred. No. 5.1e+03;		
Matches	6: Conservative	4: Mismatches	2: Indels	

QY 1 EPNHLSKIAFKI 13
| | | | | : | : | : | : | :
Db 8 ELNNLLS-VAYKV 19

RESULT 185

hydroxymethylglutaryl-CoA reductase (NADPH2) (BC 1.1.1.34) - rat (fragments)
S12554
C-Species: Rattus norvegicus (Norway rat)
C-Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C-Accession: S12554
R-Clarke, P. R.; Hardie, D.G.
EMBO J. 9, 2439-2446, 1990

A;Title: Regulation of HMG-CoA reductase: identification of the site phosphorylated by
A;Reference number: S12554; MUID:90316098; PMID:2369897
A;Accession: S12554
A;Molecule type: protein
A;Residues: 1-12;13-34 <CIA>
C;Keywords: coenzyme A; NADP; oxidoreductase; phosphoprotein

Query Match 22.2%; Score 21.5; DB 2; Length 34;
Best Local Similarity 31.8%; Pred. NO. 6.4e+03;
Matches 7; Conservative 3; Mismatches 5; Indels

```

QY      4 HLNSKIAFK-----IVSQEP 18
      | | | | : | | |
Db      2 HNRSKINQLDTPPLAVGAQEP 23

```

RESULT 186

PS03364
protein-tyrosine-phosphatase (EC 3.1.3.48) (clone PTP31) - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 20-Mar-1998
C/Accession: PS03364
R/Gen Herkog, J.; Pals, C.E.G.M.; Jonk, L.J.C.; Kruijer, W.
Biochem. Biophys. Res. Commun. 184, 1241-1249, 1992
A/Title: Differential expression of a novel murine non-receptor protein tyrosine phosphatase
A/Reference number: JH0609; MUID:92272714; PMID:1590786
A/Accession: PS03364
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-35 <DEN>
A/Experimental source: embryonal carcinoma cell, p19 cell
C/Superfamily: protein-tyrosine-phosphatase, receptor type mu; fibronectin type III receptor-tyrosine-phosphatase homology
C/Keywords: phosphoric monoester hydrolase; tyrosine-specific phosphatase

Query Match 22.2%; Score 21.5; DB 2; Length 35;
Best Local Similarity 35.3%; Pred. No. 6.6e+03;
Matches 6; Conservative 1; Mismatches 3; Indels

QY 2 PNHLNSKIAFKIVSQEP 18
DB 13 PNHY-----IATOGP 22

RESULT 187

A60365
crystsin inhibitor - tobacco hornworm (fragment)
C:Species: Manduca sexta (tobacco hornworm)
C:Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 03-Feb-1994
C:Accession: A60365
R:Khanost, M.R.
Insect Biochem. 20, 141-147, 1990
A:Title: Isolation and characterization of four serine proteinase inhibitor
A:Reference number: A60365
A:Accession: A60365
A:Molecule type: protein
A:Residues: 1-20 <KAN>
C:Keywords: hemolymph; monomer; serine proteinase inhibitor

Query Match 21.6%; Score 21; DB 2; Length 20;
Best Local Similarity 41.7%; Pred. NO. 4.3e+03;
Matches 5; Conservative 1; Mismatches 6; Indels

Qy 2 PNHLNSKIAPKI 13
||| : |||
Db 6 PNEEXGHLAFII 17

RESULT 188

S71593
serine proteinase inhibitor, 33K - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998

C;Accession: S71593
R;Rao, C.N.; Liu, Y.Y.; Peavey, C.L.; Woodley, D.T.
Arch. Biochem. Biophys. 317, 311-314, 1995
A;Title: Novel extracellular matrix-associated serine proteinase inhibitors from human s
A;Reference number: S71592; MUID:95177668; PMID:7872799
A;Accession: S71593
A;Molecule type: protein
A;Residues: 1-20 <RAO>
A;Cross-references: UNIPROT:Q9UC86
C;Function:
A;Description: involved in turnover of connective tissues
C;Keywords: serine proteinase inhibitor

Query Match 21.6%; Score 21; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 4.3e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHNSK 8
| | | | |
Db 5 EPNNGNAE 12

RESULT 189
DIRT
dentinal fluid transport-stimulating peptide - rat
N;Alternate names: DFT-stimulating peptide
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 16-Aug-2004
C;Accession: J00001
R;Yamamoto, T.; Kobayashi, M.; Yamamoto, M.; Nomura, M.; Aonuma, S.
Chem. Pharm. Bull. 34, 3803-3811, 1986
A;Title: Isolation and amino acid sequence of dentinal fluid transport-stimulating pepti
A;Reference number: J00001; MUID:87131231; PMID:3815601
A;Accession: J00001
A;Molecule type: protein
A;Residues: 1-20 <YAM>
A;Cross-references: UNIPROT:P07448
A;Experimental source: parotid gland
C;Comment: This peptide stimulates the transport of dentinal fluid, which is important f
C;Keywords: hormone; parotid gland

Query Match 21.6%; Score 21; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 4.3e+03;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 9 IAFKIVSQEP 18
| | | | |
Db 3 IAWELQHNEP 12

RESULT 190
S47207
T-cell receptor J-alpha wNVIII.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C;Accession: S47207
R;Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A;Reference number: S40133
A;Accession: S47207
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-21 <PLA>
A;Cross-references: EMBL:X71032; NID:G506610; PIDN:CAA50349.1; PID:G510318
C;Keywords: T-cell receptor

Query Match 21.6%; Score 21; DB 2; Length 21;
Best Local Similarity 33.3%; Pred. No. 4.6e+03;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SKIAFKIVSQEP 18
| | | | |
Db 2 SRVLVRLSLVP 13

RESULT 191
D30609
IG kappa chain V-III regions (Jon and Mit) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 09-Jul-2004
C;Accession: D30609
R;Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; Sol
J. Immunol. 142, 3158-3163, 1989
A;Title: Structural and idiotypic characterization of the L chains of human IgM autoant
A;Reference number: A30601; MUID:89215279; PMID:2496160
A;Accession: D30609
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-22 <GON>
A;Cross-references: UNIPROT:Q9UL83; UNIPROT:Q9UL85
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 21.6%; Score 21; DB 2; Length 22;
Best Local Similarity 42.9%; Pred. No. 4.8e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 13 IVSQEPA 19
| | | | |
Db 3 VMTQSPA 9

RESULT 192
B60691
Phycobilisome 29K linker protein - Synechococcus sp. (PCC 7002) (fragment)
C;Species: Synechococcus sp.
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 11-Jan-2000
C;Accession: B60691
R;Bryant, D.A.; de Lorimier, R.; Guglielmi, G.; Stevens Jr., S.E.
Arch. Microbiol. 153, 550-560, 1990
A;Title: Structural and compositional analyses of the phycobilisomes of Synechococcus
S.
A;Reference number: A60691; MUID:90314662; PMID:2164365
A;Accession: B60691
A;Molecule type: protein
A;Residues: 1-23 <BRV>
C;Comment: This protein, one of the eleven components detected in this species of the p
C;Superfamily: phycocyanin linker protein cpCh3
C;Keywords: Photosystem II

Query Match 21.6%; Score 21; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 5e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PNHLSKIA 10
| | | | |
Db 10 PSSQNTQVA 18

RESULT 193
A37476
hypothetical protein E3, 10.3K - human adenovirus 11 (fragment)
C;Species: Mastadenovirus h11 (human adenovirus 11)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A37476
R;Mei, Y.F.; Wadell, G.
Virology 194, 453-462, 1993
A;Title: Hemagglutination properties and nucleotide sequence analysis of the fiber gene
A;Reference number: A37476; MUID:93276532; PMID:8503168
A;Contents: Ad1p, Slobitski
A;Accession: A37476
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-26 <MEI>
A;Cross-references: UNIPROT:Q67730; UNIPROT:Q91PK8
A;Note: sequence extracted from NCBI backbone (NCBIN:132883, NCBI:P:132884)

C;Accession: S71593
R;Rao, C.N.; Liu, Y.Y.; Peavey, C.L.; Woodley, D.T.
Arch. Biochem. Biophys. 317, 311-314, 1995
A;Title: Novel extracellular matrix-associated serine proteinase inhibitors from human s
A;Reference number: S71592; MUID:95177668; PMID:7872799
A;Accession: S71593
A;Molecule type: protein
A;Residues: 1-20 <RAO>
A;Cross-references: UNIPROT:Q9UC86
C;Function:
A;Description: involved in turnover of connective tissues
C;Keywords: serine proteinase inhibitor

Query Match 21.6%; Score 21; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 4.3e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHNSK 8
| | | | |
Db 5 EPNNGNAE 12

RESULT 189
DIRT
dentinal fluid transport-stimulating peptide - rat
N;Alternate names: DFT-stimulating peptide
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 16-Aug-2004
C;Accession: J00001
R;Yamamoto, T.; Kobayashi, M.; Yamamoto, M.; Nomura, M.; Aonuma, S.
Chem. Pharm. Bull. 34, 3803-3811, 1986
A;Title: Isolation and amino acid sequence of dentinal fluid transport-stimulating pepti
A;Reference number: J00001; MUID:87131231; PMID:3815601
A;Accession: J00001
A;Molecule type: protein
A;Residues: 1-20 <YAM>
A;Cross-references: UNIPROT:P07448
A;Experimental source: parotid gland
C;Comment: This peptide stimulates the transport of dentinal fluid, which is important f
C;Keywords: hormone; parotid gland

Query Match 21.6%; Score 21; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 4.3e+03;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 9 IAFKIVSQEP 18
| | | | |
Db 3 IAWELQHNEP 12

RESULT 190
S47207
T-cell receptor J-alpha wNVIII.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C;Accession: S47207
R;Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A;Reference number: S40133
A;Accession: S47207
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-21 <PLA>
A;Cross-references: EMBL:X71032; NID:G506610; PIDN:CAA50349.1; PID:G510318
C;Keywords: T-cell receptor

Query Match 21.6%; Score 21; DB 2; Length 21;
Best Local Similarity 33.3%; Pred. No. 4.6e+03;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SKIAFKIVSQEP 18
| | | | |
Db 2 SRVLVRLSLVP 13

C;Superfamily: adenovirus early E3 10.3K protein

Query Match 21.6%; Score 21; DB 2; Length 26;
Best Local Similarity 44.4%; Pred. No. 5.8e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PNHLSKIA 10
| : | | |
Db 12 PEYRNQIA 20

RESULT 194

A45140

fatty-acyl-ethyl-ester synthase (EC 3.1.1.67) - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: A45140

R;Tsujita, T.; Okuda, H.

J. Biol. Chem. 267, 23489-23494, 1992

A;Title: Fatty acid ethyl ester synthase in rat adipose tissue and its relationship to

A;Reference number: A45140; MUID:93054696; PMID:1429692

A;Accession: A45140

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-27 <TSU>

A;Cross-references: UNIPROT:Q9R135; UNIPROT:Q91YG2

C;Superfamily: cholinesterase; cholinesterase homology

C;Keywords: carboxylic ester hydrolase

Query Match 21.6%; Score 21; DB 2; Length 27;
Best Local Similarity 33.3%; Pred. No. 6e+03;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSOE 17
| : | | |
Db 9 NTVGKVLGYVNL 23

RESULT 195

C83969

hypothetical protein BH2555 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C;Accession: C83969

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: C83969

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-28 <STO>

A;Cross-references: UNIPROT:Q9K9T0; GB:AP001515; GB:BA000004; NID:gl0174886; PIDN:BAR062

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH2555

Query Match 21.6%; Score 21; DB 2; Length 28;
Best Local Similarity 60.0%; Pred. No. 6.3e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 14 VSQEP 18
| : | | |
Db 18 ISQOP 22

RESULT 196

F97000

hypothetical protein CAC0817 [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C;Accession: F97000

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: F97000

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-28 <KUR>

A;Cross-references: UNIPROT:Q97KU9; GB:AE001437; PIDN:AAK78793.1; PID:gl5023707; GSPDB:

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC0817

Query Match 21.6%; Score 21; DB 2; Length 28;
Best Local Similarity 50.0%; Pred. No. 6.3e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 KIAFKI 13
| : | | |
Db 3 KLTFKV 8

RESULT 197

A35891

carcinoembryonic antigen-binding protein - human (fragments)

C;Species: Homo sapiens (man)

C;Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 24-Jun-1993

C;Accession: A35891

R;Toth, C.A.; Steele Jr., G.; Thomas, P.

Biochem. Biophys. Res. Commun. 171, 633-640, 1990

A;Title: A carcinoembryonic antigen (CEA) binding protein from ascites influences CEA up

A;Reference number: A35891; MUID:90386633; PMID:2403353

A;Accession: A35891

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-29 <TOT>

Query Match 21.6%; Score 21; DB 2; Length 29;
Best Local Similarity 36.4%; Pred. No. 6.5e+03;
Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKI 13
| : | | |
Db 10 NVAEXKLAEDV 20

RESULT 198

A44913

34K core flagella protein - Leptospira interrogans (fragment)

C;Species: Leptospira interrogans

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004

C;Accession: A44913; A41210

R;Trubba, G.A.; Bolin, C.A.; Zuerner, R.L.

J. Bacteriol. 174, 4761-4768, 1992

A;Title: Characterization of the periplasmic flagellum proteins of Leptospira interrogans

A;Reference number: A44913; MUID:92325069; PMID:1624463

A;Accession: A44913

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-30 <TRU>

A;Cross-references: UNIPROT:Q9R5K3

A;Experimental source: sv. pomona type kennewicki

A;Note: sequence extracted from NCBI backbone (NCBIF:108220)

C;Superfamily: flagellin

Query Match 21.6%; Score 21; DB 2; Length 30;
Best Local Similarity 33.3%; Pred. No. 6.7e+03;
Matches 7; Conservative 2; Mismatches 8; Indels 4; Gaps 1;

QY 3 NH----LNSKIAFKIVSQEPA 19
| : | | |
Db 4 NHNLAAINSHRVLFQNNVEA 24

RESULT 199		A;Status: preliminary; translated from GB/EMBL/DBJ	
B95020		A;Molecule type: DNA	
hypothetical protein SP0174 [imported] - Streptococcus pneumoniae (strain TIGR4)		A;Residues: 1-31 <RES>	
C;Species: Streptococcus pneumoniae		A;Cross-references: GB:L02326; NID:G292400; PIDN:AAA16174.1; PID:G451281	
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004		C;Genetics:	
C;Accession: B95020		A;Gene: GDB:IGLL2	
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid		A;Cross-references: GDB:137325	
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap		A;Map Position: 22q11.2-22q11.2	
nson, T.; Hickey, E.K.; Holt, I.E.		C;Superfamily: immunoglobulin C region; immunoglobulin homology	
Science 293, 498-506, 2001		Query Match 21.6%; Score 21; DB 2; Length 31;	
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,		Best Local Similarity 33.3%; Pred. No. 7e+03;	
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.		Matches 6; Conservative 2; Mismatches 10; Indels 0; Gaps 0;	
A;Reference number: A95000; MUID:21357209; PMID:11463916		Qy 1 EPNLNSKIAFKIVSQEP 18	
A;Accession: B95020		Db 2 QPXTTSPVILFLPSCEEP 19	
A;Status: preliminary		RESULT 202	
A;Molecule type: DNA		S26232	
A;Residues: 1-30 <KUR>		ribosomal protein S12 - spinach chloroplast (fragment)	
A;Cross-references: UNIPROT:Q97SY9; GB:AE005672; PIDN:AAK74355.1; PID:gl4971641; GSPDB:G		C;Species: Chloroplast Spinacia oleracea (spinach)	
A;Experimental source: strain TIGR4		C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 12-Jun-2003	
C;Genetics:		C;Accession: S26232	
A;Gene: SP0174		R;Schmidt, J.; Herfurth, E.; Subramanian, A.R.	
Query Match 21.6%; Score 21; DB 2; Length 30;		Plant Mol. Biol. 20, 459-465, 1992	
Best Local Similarity 33.3%; Pred. No. 6.7e+03;		A;Title: Purification and characterization of seven chloroplast ribosomal proteins: evi	
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;		le pathways in chloroplasts.	
Qy 5 LNSKIAFKIVSQ 16		A;Reference number: S26228; MUID:93043036; PMID:1421149	
Db 15 ISSKHCFFITDK 26		A;Accession: S26232	
RESULT 200		A;Molecule type: protein	
A72205		A;Residues: 1-31 <SCH>	
hypothetical protein - Thermotoga maritima (strain MSB8)		C;Genetics:	
C;Species: Thermotoga maritima		A;Gene: rps12	
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004		A;Genome: chloroplast	
C;Accession: A72205		C;Superfamily: ribosomal protein S12	
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey		C;Keywords: chloroplast; protein biosynthesis; ribosome	
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.		Query Match 21.6%; Score 21; DB 2; Length 31;	
C.M.		Best Local Similarity 28.6%; Pred. No. 7e+03; 5; Indels 0; Gaps 0;	
Nature 399, 323-329, 1999		Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;	
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq		Qy 6 NSKIAFKIVSQEPA 19	
A;Reference number: A72200; MUID:99287316; PMID:10360571		Db 9 NTRQPIRNVTKSPA 22	
A;Accession: A72205		RESULT 203	
A;Status: preliminary		S33176	
A;Molecule type: DNA		microtubule-associated protein MAP2C - rat (fragment)	
A;Residues: 1-30 <ARN>		C;Species: Rattus norvegicus (Norway rat)	
A;Cross-references: UNIPROT:Q9X2E9; GB:AE001820; GB:AE000512; NID:G4982411; PIDN:AAD3689		C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004	
A;Experimental source: strain MSB8		C;Accession: I55502; S33176	
C;Genetics:		R;Doll, T.; Meichsner, M.; Riederer, B.M.; Honegger, P.; Matus, A.	
A;Gene: TM1829		J. Cell Sci. 106, 663-670, 1993	
Query Match 21.6%; Score 21; DB 2; Length 30;		A;Title: An isoform of microtubule-associated protein 2 (MAP2) containing 4 repeats of	
Best Local Similarity 42.9%; Pred. No. 6.7e+03;		A;Reference number: I55502; MUID:94110306; PMID:8282771	
Matches 6; Conservative 4; Mismatches 2; Indels 1;		A;Accession: I55502	
Qy 6 NSKIAFKIVSQEPA 19		A;Status: preliminary; translated from GB/EMBL/DBJ	
Db 16 NSRPAVK--AREPS 27		A;Molecule type: mRNA	
RESULT 201		A;Residues: 1-31 <RES>	
I54515		A;Cross-references: UNIPROT:P15146; EMBL:X71487; NID:G296926; PIDN:CAAS0588.1; PID:G9393	
pre-B cell Ig lambda-like omega light chain (non-rearranging) IGLL2 - human (fragment)		C;Genetics:	
C;Species: Homo sapiens (man)		A;Gene: MAP2	
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 23-Jul-1999		C;Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology	
C;Accession: I54515		Query Match 21.6%; Score 21; DB 2; Length 31;	
R;Bauer, T.R.; McDermid, H.E.; Budarf, M.L.; Van Keuren, M.L.; Blomberg, B.B.		Best Local Similarity 50.0%; Pred. No. 7e+03;	
Immunogenetics 38, 387-399, 1993		Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;	
A;Title: Physical location of the human immunoglobulin lambda-like genes, 14.1, 16.1, an		Qy 5 LNSKIAFKIV 14	
A;Reference number: I54515; MUID:94011089; PMID:8406611		Db 5 LNSKIAFKIV 14	
A;Accession: I54515			

Db 4 LNKKMDPSKV 13

RESULT 204

A60122
fimbrial protein - Bordetella pertussis (fragment)
C/Species: Bordetella pertussis
C/Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 17-Mar-1999
C/Accession: A60122; B61474
R/Cowell, J.L.; Zhang, J.M.; Urisu, A.; Suzuki, A.; Steven, A.C.; Liu, T.; Liu, T.Y.; Ma Infect. Immun. 55, 916-922, 1987
A/Title: Purification and characterization of serotype 6 fimbriae from Bordetella pertussis
A/Reference number: A60122; MUID:87164505; PMID:2881893
A/Accession: A60122
A/Molecule type: protein
A/Residues: 1-31 <COW>
A/Experimental source: strain 114, serotype 6
R/Moibi, F.R.; van der Heide, H.G.J.; ter Avest, A.R.; Welinder, K.G.; Livey, I.; van der Microb. Pathog. 2, 473-484, 1987
A/Title: Characterization of fimbrial subunits from Bordetella species.
A/Reference number: A61474; MUID:89180964; PMID:2907088
A/Accession: B61474
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-20 <MOO>
A/Experimental source: strain BP11, serotype 3
A/Note: one strain of Bordetella pertussis may express several fimbrial proteins at once
C/Superfamily: type 1 fimbrial protein

Query Match 21.6%; Score 21; DB 2; Length 31;
Best Local Similarity 56.7%; Pred. No. 7e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EPNHLN 6
|||:
Db 20 EPSTLN 25

RESULT 205

E95140
hypothetical protein SP1211 [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C/Accession: E95140
R/Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid nson, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: E95140
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-31 <KUR>
A/Cross-references: UNIPROT:Q97QJ4; GB:AE005672; PIDN:AAK75318.1; PID:g14972692; GSPDB:G
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP1211

Query Match 21.6%; Score 21; DB 2; Length 31;
Best Local Similarity 57.1%; Pred. No. 7e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 11 FKIVSOE 17
|||:
Db 4 FKILSDX 10

RESULT 206

D70223
hypothetical protein BBD16 - Lyme disease spirochete plasmid D/lp17
C/Species: Borrelia burgdorferi (Lyme disease spirochete)

C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C/Accession: D70223
R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997
A/Authors: Smith, H.O.; Venter, J.C.
A/Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A/Reference number: A70100; MUID:98065943; PMID:9403685
A/Accession: D70223
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-31 <KLE>
A/Cross-references: UNIPROT:O51006; GB:AE000793; NID:g2689927; PIDN:AAAC66361.1; PID:g2689927
A/Experimental source: strain B31
C/Genetics:
A/Genome: plasmid

Query Match 21.6%; Score 21; DB 2; Length 31;
Best Local Similarity 44.4%; Pred. No. 7e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 LNSKIAFKI 13
|||:
Db 22 LNFVVFLI 30

RESULT 207

T06854
photosystem II protein T - Cyanophora paradoxa cyanelle
C/Species: cyanelle Cyanophora paradoxa
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C/Accession: T06854
R/Stirwalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A. submitted to the EMBL Data Library, July 1995
A/Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A/Reference number: Z15840
A/Accession: T06854
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-31 <STI>
A/Cross-references: UNIPROT:P48109; EMBL:U30821; NID:g1016083; PIDN:AAA81197.1; PID:g1016083
A/Experimental source: strain Pringsheim LB555
C/Genetics:
A/Genome: cyanelle
A/Note: psbt
C/Keywords: cyanelle; membrane-associated complex; photosynthesis; photosystem II; thyl

Query Match 21.6%; Score 21; DB 2; Length 31;
Best Local Similarity 37.5%; Pred. No. 7e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 11 FKIVSOEP 18
|||:
Db 19 FSIIFRDP 26

RESULT 208

S73244

photosystem II protein T - red alga (Porphyra purpurea) chloroplast
C/Species: chloroplast Porphyra purpurea
C/Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C/Accession: S73244
R/Reith, M.; Munholland, J. Plant Mol. Biol. Rep. 13, 333-335, 1995
A/Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.
A/Reference number: S73108
A/Accession: S73244
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-31 <REI>
A/Cross-references: UNIPROT:P51323; EMBL:U38804; NID:g1276652; PID:g1276789
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C;Genetics:
A;Gene: psbT
A;Genome: chloroplast
C;Keywords: chloroplast; membrane-associated complex; photosynthesis; photosystem II; th

Query Match 21.6%; Score 21; DB 2; Length 31;
Best Local Similarity 50.0%; Pred. No. 7e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
| | | | |
Db 17 IFFAIFREP 26

RESULT 209
F82860
hypothetical protein XF0008 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: F82860
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: F82860
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-31 <SIM>
A;Cross-references: UNIPROT:Q9PHD6; GB:AE003855; GB:AE003849; NID:g9104760; PIDN:AAF8282
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.B.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0008

Query Match 21.6%; Score 21; DB 2; Length 31;
Best Local Similarity 23.1%; Pred. No. 7e+03;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 NHLNKTIAFKIVS 15
: | : : :
Db 3 SHPSVLLMRVIS 15

RESULT 210
T07276
photosystem II protein psbT - Chlorella vulgaris chloroplast
C;Species: chloroplast Chlorella vulgaris
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C;Accession: T07276
R;Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakae
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A;Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo
A;Reference number: Z15985; MUID:97303241; PMID:9159184
A;Accession: T07276
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-31 <WAK>
A;Cross-references: UNIPROT:P56327; EMBL:AB001684; NID:g2224352; PIDN:BAA57924.1; PID:g2
C;Genetics:

A;Gene: psbT
A;Genome: chloroplast
C;Keywords: chloroplast; membrane-associated complex; photosynthesis; photosystem II; th

Query Match 21.6%; Score 21; DB 2; Length 31;
Best Local Similarity 50.0%; Pred. No. 7e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
| | | | |
Db 17 IFFAIFREP 26

RESULT 211
S39628
probable urease (EC 3.5.1.5) 69K chain - Bordetella bronchiseptica (fragment)
N;Alternate names: probable urease alpha chain
C;Species: Bordetella bronchiseptica
C;Date: 08-Jun-1994 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
C;Accession: S39628
R;Monack, D.M.; Falkow, S.
Mol. Microbiol. 10, 545-553, 1993
A;Title: Cloning of Bordetella bronchiseptica urease genes and analysis of colonization
A;Reference number: S39628; MUID:95058197; PMID:7968532
A;Accession: S39628
A;Molecule type: DNA
A;Residues: 1-32 <MON>
A;Cross-references: UNIPROT:Q9H5X1
C;Genetics:
A;Gene: ureC
C;Complex: heterodecamer; dimer of heteropentamers; heteropentamer of 1 alpha, 2 beta, 2
C;Superfamily: urease, alpha subunit; urease 62K chain homology
C;Keywords: heterodecamer; hydrolase; metalloprotein; nickel

Query Match 21.6%; Score 21; DB 2; Length 32;
Best Local Similarity 50.0%; Pred. No. 7.2e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 NHLNKSIA 10
: | : : :
Db 22 HHLDPSSIA 29

RESULT 212
S45670
Nef protein - human immunodeficiency virus type 1 (fragments)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 10-Dec-1994 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C;Accession: S45670
R;Freund, J.; Kellner, R.; Konvalinka, J.; Wolber, V.; Kraeusslich, H.G.; Kalbitzer, H.
Eur. J. Biochem. 223, 589-593, 1994
A;Title: A possible regulation of negative factor (Nef) activity of human immunodeficie
A;Reference number: S45670; MUID:94333350; PMID:8055930
A;Accession: S45670
A;Molecule type: protein
A;Residues: 1-13;14-32 <FRE>
A;Cross-references: UNIPROT:Q7LZ01
A;Note: 1-Gly, 3-Lys, 6-Lys, 7-Lys, 9-Val, and 11-Gly were also found
C;Superfamily: AIDS nef protein

Query Match 21.6%; Score 21; DB 2; Length 32;
Best Local Similarity 27.3%; Pred. No. 7.2e+03;
Matches 3; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 KIAFKIVSQEP 18
: | : : :
Db 21 EVGFVTPQVP 31

RESULT 213
C70216
hypothetical protein BBA75 - Lyme disease spirochete plasmid A/lps4
C;Species: Borrelia burgdorferi (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 C;Accession: C70216
 R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A;Authors: Smith, H.O.; Venter, J.C.
 A;Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.
 A;Reference number: A70100; MUID:98065943; PMID:9403685
 C;Accession: C70216
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-32 <KLE>
 A;Cross-references: UNIPROT:O50964; GB:AE000790; NID:G2690224; PIDN:AAC66291.1; PID:G269
 A;Experimental source: strain B31
 C;Genetics:
 A;Genome: plasmid

Query Match 21.6%; Score 21; DB 2; Length 32;
 Best Local Similarity 33.3%; Pred. No. 7.2e+03;
 Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIV 14
 ||| : : :
 Db 20 NSLNREINIDLI 31

RESULT 214
 E85588
 hypothetical protein Z1018 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL9
 C;Species: *Escherichia coli*
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C;Accession: E85588
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 C;Accession: E85588
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-32 <STO>
 A;Cross-references: UNIPROT:Q8X3V6; GB:AE005174; NID:G12513790; PIDN:AAG55169.1; GSPDB:G
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: Z1018

Query Match 21.6%; Score 21; DB 2; Length 32;
 Best Local Similarity 57.1%; Pred. No. 7.2e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 SKIAFKI 13
 ||| : : :
 Db 21 SKWAMKL 27

RESULT 215
 S08482
 regulatory protein LAC9-2 - yeast (*Kluyveromyces marxianus* var. *lactis*) (fragment)
 N;Alternate names: transcription activator protein LAC9
 C;Species: *Kluyveromyces marxianus* var. *lactis*, *Candida sphaerica*
 C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Aug-2004
 C;Accession: S08482
 R;Kuger, P.; Goedecke, A.; Breunig, K.D.
 Nucleic Acids Res. 18, 745-751, 1990
 A;Title: A mutation in the Zn-finger of the GAL4 homolog LAC9 results in glucose repress
 A;Reference number: S08482; MUID:90192144; PMID:2107531
 C;Accession: S08482
 A;Molecule type: DNA
 A;Residues: 1-32 <KUG>
 A;Cross-references: UNIPROT:P08657; EMBL:X17249
 A;Note: the authors translated the codon CAG for residue 19 as Glu and TCG for residue 2
 C;Genetics:

A;Gene: LAC9-2
 C;Superfamily: GAL4 zinc binuclear cluster homology
 C;Keywords: DNA binding; nucleus; transcription regulation

Query Match 21.6%; Score 21; DB 2; Length 32;
 Best Local Similarity 28.6%; Pred. No. 7.2e+03;
 Matches 4; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQ 16
 ||| : : :
 Db 6 NNNNNKKSSEVMHQ 19

RESULT 216
 AF0546
 hypothetical protein STY0391 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typh
 C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A;Note: this species has also been called *Salmonella typhi*
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AF0546
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* sero
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 C;Accession: AF0546
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-33 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD08814.1; PID:G16501628; GSPDB:GN00176
 C;Genetics:
 A;Gene: STY0391

Query Match 21.6%; Score 21; DB 2; Length 33;
 Best Local Similarity 26.7%; Pred. No. 7.5e+03;
 Matches 4; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
 ||| : : :
 Db 2 LNRYRFAVEKDDPS 16

RESULT 217
 D70224
 hypothetical protein BBE01 - Lyme disease spirochaete plasmid E/lp25
 C;Species: *Borrelia burgdorferi* (Lyme disease spirochaete)
 C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 C;Accession: D70224
 R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A;Authors: Smith, H.O.; Venter, J.C.
 A;Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.
 A;Reference number: A70100; MUID:98065943; PMID:9403685
 C;Accession: D70224
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-33 <KLE>
 A;Cross-references: UNIPROT:O50697; GB:AE000785; NID:G2689951; PIDN:AAC66047.1; PID:G268
 A;Experimental source: strain B31
 C;Genetics:
 A;Genome: plasmid

Query Match 21.6%; Score 21; DB 2; Length 33;
 Best Local Similarity 13.3%; Pred. No. 7.5e+03;
 Matches 2; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQE 17
 ||| : : :
 Db 8 DNLDKEVVFETKD 22

RESULT 218

E82303
hypothetical protein VC0588 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: E82303

R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: E82303

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-33 <HEI>

A:Cross-references: UNIPROT:Q9KUD4; GB:AE004144; GB:AE003852; NID:g9655017; PIDN:AAF9375

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC0588

A:Map position: 1

Query Match 21.6%; Score 21; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNH 4

|||

22 PNH 24

RESULT 219

E82391

hypothetical protein VCA0995 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: E82391

R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: E82391

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-33 <HEI>

A:Cross-references: UNIPROT:Q9KKV5; GB:AE004426; GB:AE003853; NID:g9658431; PIDN:AAF9689

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0995

A:Map position: 2

Query Match 21.6%; Score 21; DB 2; Length 33;
Best Local Similarity 40.0%; Pred. No. 7.5e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EPNHLNSKIA 10

|||

22 KPNDLALQVA 31

RESULT 220

S58578

hypothetical protein 33 - maize chloroplast

C:Species: chloroplast Zea mays (maize)

C>Date: 29-Nov-1995 #sequence_revision 19-Jan-1996 #text_change 09-Jul-2004

C:Accession: S58578

R:Maier, R.M.; Neckeremann, K.; Igloi, G.L.; Koessel, H.

J. Mol. Biol. 251, 614-628, 1995

A>Title: Complete sequence of the maize chloroplast genome: gene content, hotspots of di

A:Reference number: S58531; MUID:95395841; PMID:7666415

A:Accession: S58578

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-33 <MAI>

A:Cross-references: UNIPROT:P37257; EMBL:X86563; NID:g902200; PIDN:CAAG0312.1; PID:g902202

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

C:Genetics:

A:Genome: chloroplast

C:Keywords: chloroplast

Query Match 21.6%; Score 21; DB 2; Length 33;

Best Local Similarity 50.0%; Pred. No. 7.5e+03;

Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 IAFKIVSQEP 18

|||

17 IFFAIFREP 26

RESULT 221

A44038

leader peptide TnaC - Proteus vulgaris

C:Species: Proteus vulgaris

C>Date: 10-Jun-1993 #sequence_revision 19-Nov-1994 #text_change 09-Jul-2004

C:Accession: A44038

R:Kamath, A.V.; Yanofsky, C.

J. Biol. Chem. 267, 19978-19985, 1992

A>Title: Characterization of the tryptophanase operon of Proteus vulgaris. Cloning, nucle

A:Reference number: A44038; MUID:93015858; PMID:1400314

A:Accession: A44038

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-34 <KAM>

A:Cross-references: UNIPROT:P28779; GB:M93277; NID:g150910; PIDN:AAA25663.1; PID:g15091

A>Note: sequence extracted from NCBI backbone (NCBIP:115636)

C:Superfamily: unassigned leader peptides

Query Match 21.6%; Score 21; DB 2; Length 34;

Best Local Similarity 37.5%; Pred. No. 7.7e+03;

Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLNSKIAP 11

|||

22 NIDSELAF 29

RESULT 222

C95037

hypothetical protein SP0316 [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C:Accession: C95037

R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei

son, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison

A>Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: C95037

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-35 <KUR>

A:Cross-references: UNIPROT:Q97SL6; GB:AE005672; PIDN:AAK74492.1; PID:gl4971789; GSPDB

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0316

Query Match 21.6%; Score 21; DB 2; Length 35;

Best Local Similarity 100.0%; Pred. No. 8e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

16087Y

pancreatic hormone - bullfrog		pancreatic polypeptide		pancreatic polypeptide	
Alternate names: pancreatic polypeptide		Alternate names: pancreatic polypeptide		Alternate names: pancreatic polypeptide	
Species: Rana catesbeiana (bullfrog)		Species: Rana catesbeiana (bullfrog)		Species: Rana catesbeiana (bullfrog)	
Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004		Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004		Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004	
Accession: A28091		Accession: A28091		Accession: A28091	
Pollock, H.G.; Hamilton, J.W.; Rouse, J.B.; Ebner, K.E.; Rawitch, A.B.		Pollock, H.G.; Hamilton, J.W.; Rouse, J.B.; Ebner, K.E.; Rawitch, A.B.		Pollock, H.G.; Hamilton, J.W.; Rouse, J.B.; Ebner, K.E.; Rawitch, A.B.	
Biol. Chem. 263, 9746-9751, 1988		Biol. Chem. 263, 9746-9751, 1988		Biol. Chem. 263, 9746-9751, 1988	
Title: Isolation of peptide hormones from the pancreas of the bullfrog (Rana catesbeiana)		Title: Isolation of peptide hormones from the pancreas of the bullfrog (Rana catesbeiana)		Title: Isolation of peptide hormones from the pancreas of the bullfrog (Rana catesbeiana)	
Reference number: A92730; MUID:88257102; PMID:3260236		Reference number: A92730; MUID:88257102; PMID:3260236		Reference number: A92730; MUID:88257102; PMID:3260236	
Accession: A28091		Accession: A28091		Accession: A28091	
Molecule type: protein		Molecule type: protein		Molecule type: protein	
Residues: 1-36 <POL>		Residues: 1-36 <POL>		Residues: 1-36 <POL>	
Cross-references: UNIPROT:P15427		Cross-references: UNIPROT:P15427		Cross-references: UNIPROT:P15427	
Superfamily: pancreatic hormone		Superfamily: pancreatic hormone		Superfamily: pancreatic hormone	
Query Match 21.6%; Score 21; DB 2; Length 36;		Query Match 21.6%; Score 21; DB 2; Length 36;		Query Match 21.6%; Score 21; DB 2; Length 36;	
Best Local Similarity 75.0%; Pred. No. 8.2e+03;		Best Local Similarity 75.0%; Pred. No. 8.2e+03;		Best Local Similarity 75.0%; Pred. No. 8.2e+03;	
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY 1 EPNH 4		QY 1 EPNH 4		QY 1 EPNH 4	
: :		: :		: :	
DB 4 EPHH 7		DB 4 EPHH 7		DB 4 EPHH 7	
RESULT 228		RESULT 228		RESULT 228	
E69729		E69729		E69729	
Required for translation of spoIIID used - Bacillus subtilis		Required for translation of spoIIID used - Bacillus subtilis		Required for translation of spoIIID used - Bacillus subtilis	
Species: Bacillus subtilis		Species: Bacillus subtilis		Species: Bacillus subtilis	
Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004		Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004		Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004	
Accession: E69729		Accession: E69729		Accession: E69729	
Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bester, C.; Bron, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.		Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bester, C.; Bron, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.		Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bester, C.; Bron, S.; Brouillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.	
Nature 390, 249-256, 1997		Nature 390, 249-256, 1997		Nature 390, 249-256, 1997	
Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.		Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.		Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.	
Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.		Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.		Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.	
Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.		Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.		Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.	
Reference number: A69580; MUID:98044033; PMID:9384377		Reference number: A69580; MUID:98044033; PMID:9384377		Reference number: A69580; MUID:98044033; PMID:9384377	
Accession: E69729		Accession: E69729		Accession: E69729	
Status: preliminary; nucleic acid sequence not shown; translation not shown		Status: preliminary; nucleic acid sequence not shown; translation not shown		Status: preliminary; nucleic acid sequence not shown; translation not shown	
Molecule type: DNA		Molecule type: DNA		Molecule type: DNA	
Residues: 1-36 <KUN>		Residues: 1-36 <KUN>		Residues: 1-36 <KUN>	
Cross-references: UNIPROT:Q32276; GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15660.		Cross-references: UNIPROT:Q32276; GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15660.		Cross-references: UNIPROT:Q32276; GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15660.	
Experimental source: strain 168		Experimental source: strain 168		Experimental source: strain 168	
Genetics:		Genetics:		Genetics:	
Gene: usd		Gene: usd		Gene: usd	
Query Match 21.6%; Score 21; DB 2; Length 36;		Query Match 21.6%; Score 21; DB 2; Length 36;		Query Match 21.6%; Score 21; DB 2; Length 36;	
Best Local Similarity 33.3%; Pred. No. 8.2e+03;		Best Local Similarity 33.3%; Pred. No. 8.2e+03;		Best Local Similarity 33.3%; Pred. No. 8.2e+03;	
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	
QY 8 KIAFKIVSQ 16		QY 8 KIAFKIVSQ 16		QY 8 KIAFKIVSQ 16	
: :		: :		: :	
DB 6 KLTFTKIQR 14		DB 6 KLTFTKIQR 14		DB 6 KLTFTKIQR 14	
RESULT 229		RESULT 229		RESULT 229	
A69287		A69287		A69287	
Hypothetical protein AF0297 - Archaeoglobus fulgidus		Hypothetical protein AF0297 - Archaeoglobus fulgidus		Hypothetical protein AF0297 - Archaeoglobus fulgidus	
Species: Archaeoglobus fulgidus		Species: Archaeoglobus fulgidus		Species: Archaeoglobus fulgidus	
Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004		Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004		Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004	
Accession: A69287		Accession: A69287		Accession: A69287	
Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.		Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.		Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.	


```

RESULT 232
T07292
hypothetical protein ycf12 - Chlorella vulgaris chloroplast
C:Species: chloroplast Chlorella vulgaris
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C:Accession: T07292
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakas
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo
A:Reference number: Z15985; MUID:97303241; PMID:9159184
A:Accession: T07292
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-37 <WAK>
A:Cross-references: UNIPROT:P56328; EMBL:AB001684; NID:g2224352; PIDN:BAA57940.1; PID:g2
C:Genetics:
A:Gene: ycf12
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 21.6%; Score 21; DB 2; Length 37;
Best Local Similarity 27.3%; Pred. No. 8.5e+03;
Matches 3; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVS 15
:|:|:|:|:|:
Db 5 MNLEIVQLTA 15

RESULT 233
AH0844
hypothetical protein STV2958 [imported] - Salmonella enterica subsp. enterica serovar Ty
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0844
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH0844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-37 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05943.1; PID:g16503914; GSPDB:GN00176
C:Genetics:
A:Gene: STV2958

Query Match 21.6%; Score 21; DB 2; Length 37;
Best Local Similarity 22.2%; Pred. No. 8.5e+03;
Matches 2; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EPNHLNSKI 9
:|:|:|:|:|:
Db 24 KPTHMQREV 32

RESULT 234
F2KTL
photosystem II protein psbL - Cyanophora paradoxa cyanelle
N:Alternate names: photosystem II protein psbH
C:Species: cyanelle Cyanophora paradoxa
C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: S09483; T06868
R:Cantrill, A.; Bryant, D.A.
Prog. Photosyn. Res. 4, 659-662, 1987
A:Title: Molecular cloning and nucleotide sequences of the genes encoding cytochrome B-5
A:Reference number: S09182

```

```

A:Accession: S09483
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-38 <CAN>
A:Cross-references: UNIPROT:P19154
R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.
submitted to the EMBL Data Library, July 1995
A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A:Reference number: Z15840
A:Accession: T06868
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-38 <STJ>
A:Cross-references: EMBL:U30821; NID:g1016083; PIDN:AAA81211.1; PID:g1016124
A:Experimental source: strain Pringsheim LB555
C:Genetics:
A:Gene: psbL; psbH
A:Genome: cyanelle
C:Superfamily: photosystem II protein psbL
C:Keywords: cyanelle; photosynthesis; Photosystem II; thylakoid; transmembrane protein
F:18-37/Domain: transmembrane #status predicted <TMM>

Query Match 21.6%; Score 21; DB 1; Length 38;
Best Local Similarity 56.7%; Pred. No. 8.7e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 IVSQEP 18
:|:|:|:|:|:
Db 1 MVSQNP 6

RESULT 235
S73157
Photosystem II protein I - red alga (Porphyra purpurea) chloroplast
C:Species: chloroplast Porphyra purpurea
C>Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: S73157
R:Reith, M.; Munnolland, J.
Plant Mol. Biol. Rep. 13, 333-335, 1995
A:Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.
A:Reference number: S73108
A:Accession: S73157
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-38 <REI>
A:Cross-references: UNIPROT:P51236; EMBL:U38804; NID:g1276652; PIDN:AAC08122.1; PID:g12
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C:Genetics:
A:Gene: psbI
A:Genome: chloroplast
C:Superfamily: photosystem II protein psbI
C:Keywords: chloroplast; photosynthesis; Photosystem II; thylakoid

Query Match 21.6%; Score 21; DB 2; Length 38;
Best Local Similarity 30.8%; Pred. No. 8.7e+03;
Matches 4; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 7 SXIAFKIVSQEPA 19
:|:|:|:|:|:
Db 17 SLFVFGFLSNDFS 29

RESULT 236
S68260
hypothetical protein gadd7.1 - long-tailed hamster
C:Species: Cricetus longicaudatus (long-tailed hamster)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S68260
R:Hollander, M.C.; Alamo, I.; Fornace Jr., A.J.
Nucleic Acids Res. 24, 1589-1593, 1996
A:Title: A novel DNA damage-inducible transcript, gadd7, inhibits cell growth, but lacks
A:Reference number: S68260; MUID:96211359; PMID:8649973
A:Accession: S68260

```

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-38 <HOL>
A;Cross-references: EMBL:L40430

Query Match 21.6%; Score 21; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 8.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNH 4
|||
Db 23 PNH 25

RESULT 237

C70222

hypothetical protein BB07 - Lyme disease spirochete plasmid D/Ip17

C;Species: Borrelia burgdorferi (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C;Accession: C70222

R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
i. Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997

A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A;Reference number: A70100; MUID:98065943; PMID:9403685

A;Accession: C70222

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-38 <KLE>

A;Cross-references: UNIPROT:O51002; GB:AE000793; NID:g2689927; PIDN:AAC66364.1; PID:g268

A;Experimental source: strain B31

C;Genetics:

A;Genome: plasmid

Query Match 21.6%; Score 21; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 8.7e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAFK 12
||| : ||
Db 26 LNKPLIFK 33

RESULT 238

JN0418

hypothetical protein OPC38 (psbH region) - rye chloroplast

C;Species: chloroplast Secale cereale (rye)

C;Date: 05-Mar-1993 #sequence_revision 05-Mar-1993 #text_change 09-Jul-2004

C;Accession: JN0418

R;Bukharov, A.A.; Kolosov, V.L.; Zolotarev, A.S.; Abdulaev, N.G.

Bioorg. Khim. 15, 927-939, 1989

A;Title: Rye photosystem II: nucleotide sequence of psbH and psbH genes coding for 47kD

A;Reference number: JN0415; MUID:90056620; PMID:2684168

A;Accession: JN0418

A;Molecule type: DNA

A;Residues: 1-38 <BUK>

A;Cross-references: UNIPROT:P37260

A;Note: article in Russian with English abstract

C;Genetics:

A;Genome: chloroplast

C;Keywords: chloroplast; transmembrane protein

F;3-23/Domain: transmembrane #status predicted <TMM>

Query Match 21.6%; Score 21; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 8.7e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
||| : ||
Db 17 IFFAIFRREP 26

RESULT 239

S14141

hypothetical protein 38 - wheat chloroplast

C;Species: chloroplast Triticum aestivum (common wheat)

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C;Accession: S14141; S14563

R;Hird, S.M.; Webber, A.N.; Wilson, R.J.; Dyer, T.A.; Gray, J.C.

Curr. Genet. 19, 199-206, 1991

A;Title: Differential expression of the psbB and psbH genes encoding the 47 kDa chlorop

A;Reference number: S14140; MUID:91330334; PMID:1714358

A;Accession: S14141

A;Molecule type: DNA

A;Residues: 1-38 <HIR>

A;Cross-references: UNIPROT:P37260; EMBL:X54749; NID:g12356; PIDN:CAA38542.1; PID:g1235

C;Genetics:

A;Genome: chloroplast

C;Keywords: chloroplast; transmembrane protein

Query Match 21.6%; Score 21; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 8.7e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
||| : ||
Db 17 IFFAIFRREP 26

RESULT 240

I56173

EBV/C3d receptor - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000

C;Accession: I56173

R;Sirha, S.K.; Todd, S.C.; Hedrick, J.A.; Speiser, C.L.; Lambris, J.D.; Tsoukas, C.D.

J. Immunol. 150, 5311-5320, 1993

A;Title: Characterization of the EBV/C3d receptor on the human Jurkat T cell line: evid

A;Reference number: I56173; MUID:93294286; PMID:8390533

A;Accession: I56173

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-39 <RES>

A;Cross-references: GB:S62696; NID:g386263; PIDN:AAB27186.1; PID:g386264

Query Match 21.6%; Score 21; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 9e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHL 5
|||
Db 6 EGNHL 10

RESULT 241

C90523

hypothetical protein MYP 0910 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C;Species: Mycoplasma pulmonis

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C;Accession: C90523

R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.

Nucleic Acids Res. 29, 2145-2153, 2001

A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul

A;Reference number: A99512; MUID:21267165; PMID:11353084

A;Accession: C90523

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-39 <KUR>

A;Cross-references: UNIPROT:Q98RB8; GB:AL445566; PID:g4089504; PIDN:CAC13264.1; GSPDB

A;Experimental source: strain UAB CTIP

C;Genetics:

A;Gene: MYP 0910

A;Genetic code: SGC3

Query Match 21.6%; Score 21; DB 2; Length 39;
 Best Local Similarity 44.4%; Pred. No. 9e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NNLNSKIAF 11
 :||:|
 Db 30 SHINFKGLF 38

RESULT 242
 C72426
 hypothetical protein - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: C72426
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.N.

Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: C72426
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-39 <ARN>
 A:Cross-references: UNIPROT:Q9WXQ3; GB:AB001691; GB:AE000512; NID:g4980517; PIDN:AAD3514
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TW0046

Query Match 21.6%; Score 21; DB 2; Length 39;
 Best Local Similarity 80.0%; Pred. No. 9e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 VSQEP 18
 :||:|
 Db 19 VSKEP 23

RESULT 243
 S06901
 hypothetical protein 39 - Synechococcus sp. (PCC 7002)
 C:Species: Synechococcus sp.
 C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
 C:Accession: S06901
 R:Gantrell, A.; Bryant, D.A.
 Plant Mol. Biol. 9, 453-468, 1987
 A:Title: Molecular cloning and nucleotide sequence of the psaa and psab genes of the cya
 A:Reference number: S06397
 A:Accession: S06901
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-39 <CAN>
 A:Cross-references: UNIPROT:Q55324

Query Match 21.6%; Score 21; DB 2; Length 39;
 Best Local Similarity 57.1%; Pred. No. 9e+03;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 HLNSKIA 10
 :||:|
 Db 20 HLPSQLA 26

RESULT 244
 D83811
 hypothetical protein BH1292 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C:Accession: D83811
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: D83811
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-39 <STO>
 A:Cross-references: UNIPROT:Q9KDC1; GB:AF001511; GB:BA000004; NID:gi0173727; PIDN:BA05
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH1292

Query Match 21.6%; Score 21; DB 2; Length 39;
 Best Local Similarity 75.0%; Pred. No. 9e+03;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NNLN 6
 :||:|
 Db 36 NNVN 39

RESULT 245
 A61320
 plastocyanin - Anemone nemorosa (fragment)
 C:Species: Anemone nemorosa
 C>Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
 C:Accession: A61320
 R:Grund, C.; Gilroy, J.; Gleaves, T.; Jensen, U.; Boulter, D.
 Phytochemistry 20, 1559-1565, 1981
 A:Title: Systematic relationships of the Ranunculaceae based on amino acid sequence dat

A:Reference number: A61320
 A:Accession: A61320
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-40 <GRU>
 A:Cross-references: UNIPROT:Q7M253
 C:Superfamily: plastocyanin
 C:Keywords: electron transfer; metalloprotein

Query Match 21.6%; Score 21; DB 2; Length 40;
 Best Local Similarity 40.0%; Pred. No. 9.3e+03;
 Matches 6; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 2 PNHLN----SKIAFK 12
 :||:|
 Db 16 PMSFVAAAGEKIVFK 30

RESULT 246
 C61320
 plastocyanin - Clematis vitalba (fragment)
 C:Species: Clematis vitalba
 C>Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
 C:Accession: C61320
 R:Grund, C.; Gilroy, J.; Gleaves, T.; Jensen, U.; Boulter, D.
 Phytochemistry 20, 1559-1565, 1981
 A:Title: Systematic relationships of the Ranunculaceae based on amino acid sequence data

A:Reference number: A61320
 A:Accession: C61320
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-40 <GRU>
 A:Cross-references: UNIPROT:Q7M257
 C:Superfamily: plastocyanin
 C:Keywords: electron transfer; metalloprotein

Query Match 21.6%; Score 21; DB 2; Length 40;
 Best Local Similarity 40.0%; Pred. No. 9.3e+03;
 Matches 6; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 2 PNHLN----SKIAFK 12
 :||:|
 Db 16 PMSFVAAAGEKIVFK 30

RESULT 247

E61320
 plastocyanin - Eranthis hyemalis (fragment)
 C;Species: Eranthis hyemalis
 C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
 C;Accession: E61320
 R;Grund, C.; Gilroy, J.; Gleaves, T.; Jensen, U.; Boulter, D.
 Phytochemistry 20, 1559-1565, 1981
 A;Title: Systematic relationships of the Ranunculaceae based on amino acid sequence data
 A;Reference number: A61320
 A;Accession: E61320
 A;Status: Preliminary
 A;Molecule type: protein
 A;Residues: 1-40 <GRU>
 A;Cross-references: UNIPROT:Q7M254
 C;Superfamily: plastocyanin
 C;Keywords: electron transfer; metalloprotein

Query Match 21.6%; Score 21; DB 2; Length 40;
 Best Local Similarity 40.0%; Pred. No. 9.3e+03;
 Matches 6; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 2 PNHLN----SKIAFK 12
 ||| : |||
 Db 16 PMSFSVAPGEKIVFK 30

RESULT 248

PC4021
 chaperonin containing t-complex polypeptide 1 - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 26-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 04-Sep-1998
 C;Accession: PC4021
 R;Kubota, H.; Hynes, G.; Willison, K.
 Gene 154, 231-236, 1995
 A;Title: The eighth Cct gene, Cctq, encoding the theta subunit of the cytosolic chaperonin
 A;Reference number: JC4073; MUID:95197008; PMID:7890169
 A;Accession: PC4021
 A;Molecule type: mRNA
 A;Residues: 1-40 <KUB>
 A;Cross-references: EMBL:Z37163
 C;Genetics:
 A;Gene: Cctq
 C;Superfamily: molecular chaperone t-complex-type
 C;Keywords: molecular chaperone

Query Match 21.6%; Score 21; DB 2; Length 40;
 Best Local Similarity 37.5%; Pred. No. 9.3e+03;
 Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNHLNSKI 9
 ||| :
 Db 17 PNGMKMV 24

RESULT 249

A40128
 probable antigen 10 - Mycobacterium leprae (fragment)
 C;Species: Mycobacterium leprae
 C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 18-Jun-1993
 C;Accession: A40128
 R;Hartsteerl, R.A.; van Rens, R.M.; Stabel, L.F.E.M.; de Wit, M.Y.L.; Klatser, P.R.
 Infect. Immun. 58, 2821-2827, 1990
 A;Title: Selection and characterization of recombinant clones that produce Mycobacterium
 A;Reference number: A41476; MUID:90354041; PMID:1696931
 A;Accession: A40128
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-40 <HAR>

Query Match 21.6%; Score 21; DB 2; Length 40;
 Best Local Similarity 66.7%; Pred. No. 9.3e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 14 VSQEP 19
 : |||
 Db 12 LSQHP 17

RESULT 250

D82440
 hypothetical protein VCA0598 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C;Accession: D82440
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
 L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82440
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-40 <HEI>
 A;Cross-references: UNIPROT:Q9KLZ3; GB:A5004390; GB:A5003853; NID:9658003; PIDN:AAF964
 A;Experimental source: serogroup O1; strain N16961; biotype E1 Tor
 C;Genetics:
 A;Gene: VCA0598
 A;Map position: 2

Query Match 21.6%; Score 21; DB 2; Length 40;
 Best Local Similarity 44.4%; Pred. No. 9.3e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNHLNSKIA 10
 ||| : |||
 Db 23 PFHQSSAVA 31

Search completed: October 19, 2004, 19:14:20
 Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 19, 2004, 18:56:58 ; Search time 189 Seconds
(without alignments)
57.842 Million cell updates/sec

Title: US-10-799-005a-1
Perfect score: 97
Sequence: 1 EPNHLNKSIAFKIVSQEPA 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 67787

Minimum DB seq length: 19
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	39.2	50	2	Q7R8B1
2	36	37.1	20	2	Q6JDL7
3	35.5	36.6	41	2	Q8FFB7
4	35	36.1	21	2	Q712N7
5	35	36.1	21	2	CAB36907
6	34	35.1	38	2	O85326
7	33	34.0	20	2	Q9PSI5
8	33	34.0	41	2	O8FFB5
9	33	34.0	50	2	Q9PCZ2
10	32	33.0	35	2	Q9RHG5
11	32	33.0	37	2	Q7LYB7
12	32	33.0	41	2	Q84XN8
13	32	33.0	41	2	Q84XN9
14	32	33.0	41	2	Q84XPO
15	32	33.0	41	2	Q84XP1
16	32	33.0	41	2	Q7VGM8
17	31	32.0	31	2	Q9UDE5
18	31	32.0	31	2	Q8E1R7
19	31	32.0	43	2	Q7UEV7
20	31	32.0	50	2	O8CKB8
21	30.5	31.4	42	1	RS7 STAAU
22	30.5	31.4	45	2	Q9W6B7
23	30	30.9	31	2	Q7RM25
24	30	30.9	34	2	Q71VB2
25	30	30.9	34	2	Q8NV10
26	30	30.9	34	2	Q91Y56
27	30	30.9	34	2	AAB97515
28	30	30.9	34	2	AAB84098
29	30	30.9	36	2	O8FD47
30	30	30.9	37	2	Q7RED1
31	30	30.9	37	2	Q817S9

32	30	30.9	37	2	Q8F8V5	Q8f8v5 leptospira
33	30	30.9	42	2	Q8CR89	Q8cr89 staphylococ
34	30	30.9	43	2	Q96BN7	Q96bn7 homo sapien
35	30	30.9	46	1	Y576 PYRFU	Y8u394 pyrococcus
36	30	30.9	46	2	Q6PKM7	Q6pkm7 homo sapien
37	30	30.9	46	2	AAT02226	Aat02226 homo sapi
38	30	30.9	47	2	Q98156	Q98156 human herpe
39	29.5	30.4	30	2	Q7VKW6	Q7vkw6 haemophilus
40	29.5	30.4	32	2	Q8DY47	Q8dy47 streptococc
41	29	29.9	33	2	Q9KPS8	Q9kps8 vibrio chol
42	29	29.9	34	2	Q8F830	Q8f830 leptospira
43	29	29.9	40	2	Q9KMU3	Q9kmu3 vibrio chol
44	29	29.9	45	2	Q7QXP2	Q7qxp2 giardia lam
45	29	29.9	46	2	Q8F042	Q8f042 leptospira
46	29	29.9	47	2	Q72DQ2	Q72dq2 desulfovibr
47	29	29.9	47	2	Q7TDN5	Q7tdn5 halovirus h
48	29	29.9	47	2	Q8V6V0	Q8v6v0 halovirus h
49	29	29.9	47	2	AAS95357	Aas95357 desulfovi
50	29	29.9	48	2	Q9SN96	Q9sn96 sus scrofa
51	29	29.9	49	2	Q95418	Q95418 homo sapien
52	29	29.9	50	2	Q7YS12	Q7ys12 ovis aries
53	29	29.9	50	2	Q8XMB6	Q8xmb6 clostridium
54	28.5	29.4	49	2	Q9A828	Q9a828 caulobacter
55	28	28.9	23	2	Q9AJQ7	Q9ajq7 vibrio fisc
56	28	28.9	24	2	Q01228	Q01228 williopsis
57	28	28.9	25	1	SPIG_PSEUS	SP83357 pseudacanth
58	28	28.9	25	2	Q9TRB6	Q9tre6 cercopithe
59	28	28.9	34	2	Q97PI6	Q97pi6 streptococ
60	28	28.9	35	2	Q71M23	Q71m23 streptococ
61	28	28.9	35	2	AAQ04270	AAQ04270 streptoco
62	28	28.9	37	2	Q9TTH7	Q9tth7 bos taurus
63	28	28.9	38	2	Q7PH38	Q7ph38 anopheles g
64	28	28.9	41	2	Q8CRL7	Q8crl7 staphylococ
65	28	28.9	41	2	Q8EZK3	Q8ezk3 leptospira
66	28	28.9	43	2	Q7TDL0	Q7tdl0 halovirus h
67	28	28.9	43	2	Q8V6S1	Q8v6s1 halovirus h
68	28	28.9	46	2	Q8VB64	Q8ve64 mus musculu
69	28	28.9	49	2	Q9UGN7	Q9ugn7 haemonchus
70	28	28.9	49	2	Q9UGN8	Q9ugn8 haemonchus
71	28	28.9	49	2	Q9UGN9	Q9ugn9 haemonchus
72	28	28.9	49	2	Q7MBK7	Q7mbk7 vibrio vuln
73	27.5	28.4	28	2	Q8CLV6	Q8clv6 yersinia pe
74	27.5	28.4	30	2	Q8FZX9	Q8fzx9 brucella su
75	27.5	28.4	36	2	Q7RMJ0	Q7rmj0 plasmodium
76	27	27.8	25	2	Q9Y565	Q9y565 homo sapien
77	27	27.8	27	2	Q9UWG8	Q9uwg8 thermococcu
78	27	27.8	29	2	P78747	P78747 saccharomyc
79	27	27.8	31	2	Q8WD36	Q8wd36 melipona bi
80	27	27.8	33	1	RL26_XENLA	RL2629 xenopus lae
81	27	27.8	33	2	Q7INE3	Q7ine3 oxyrrhynch
82	27	27.8	33	2	AAQ03937	AAQ03937 oxyrrhynch
83	27	27.8	34	2	Q8WY57	Q8wy57 homo sapien
84	27	27.8	36	2	Q7M4R4	Q7m4r4 homo sapien
85	27	27.8	36	2	Q9RR54	Q9rr54 salmonella
86	27	27.8	37	2	Q878S4	Q878s4 streptococ
87	27	27.8	37	2	Q8F140	Q8f140 streptococ
88	27	27.8	38	1	Y520_BORBU	Y520 borrelia bu
89	27	27.8	38	2	Q8TGN0	Q8tgn0 saccharomyc
90	27	27.8	38	2	Q15931	Q15931 homo sapien
91	27	27.8	38	2	Q9GAU0	Q9gau0 rana sylvat
92	27	27.8	39	2	Q925D3	Q925d3 rattus norv
93	27	27.8	41	2	Q9RG94	Q9rg94 borrelia he
94	27	27.8	41	2	Q7MJ15	Q7mj15 vibrio vuln
95	27	27.8	42	2	Q7QWJ3	Q7qwj3 giardia lam
96	27	27.8	44	2	Q9GAT8	Q9gat8 rana sylvat
97	27	27.8	44	2	Q7P9P5	Q7p9p5 rickettsia
98	27	27.8	44	2	Q72AJ8	Q72aj8 desulfovibr
99	27	27.8	44	2	AAS96470	Aas96470 desulfovi
100	27	27.8	45	2	Q9S4W1	Q9s4w1 escherichia
101	27	27.8	45	2	Q83CF9	Q83cf9 coxiella bu
102	27	27.8	46	2	Q82Z68	Q82z68 enterococcu
103	27	27.8	47	2	Q9BXX5	Q9bxx5 homo sapien
104	27	27.8	47	2	Q9GAU3	Q9gau3 rana sylvat

105	27	27.8	47	2	Q8FDQ3	Q8fdq3 escherichia	178
106	27	27.8	47	2	Q9X1D2	Q9x1d2 thermotoga	179
107	27	27.8	47	2	Q6L7B1	Q6l7b1 ctenopharyn	180
108	27	27.8	48	1	PSD_AZOV1	Q44558 azotobacter	181
109	27	27.8	48	2	Q6PVS2	Q6pvs2 ovis aries	182
110	27	27.8	48	2	Q8K264	Q8ke64 chlorobium	183
111	27	27.8	48	2	Q8K264	Q8ke64 chlorobium	184
112	27	27.8	48	2	AAS87353	Aas87353 ovis arie	185
113	26.5	27.3	50	2	Q9AGM9	Q9agm9 clostridium	186
114	26.5	27.3	32	2	Q7VLJ7	Q7vlj7 haemophilus	187
115	26.5	27.3	36	2	Q9RHE3	Q9rhe3 pediococcus	188
116	26.5	27.3	37	2	Q9PC19	Q9pci9 xyella fas	189
117	26.5	27.3	39	2	Q727X8	Q727x8 desulfobivr	190
118	26.5	27.3	41	2	AAS97198	Aas97198 desulfovi	191
119	26	26.8	19	2	Q99J74	Q99jt4 mus musculu	192
120	26	26.8	21	2	Q9BX20	Q9bx20 homo sapien	193
121	26	26.8	23	2	Q9X3G5	Q9x3g5 prochloroco	194
122	26	26.8	23	2	Q9LA84	Q9la84 rhodobacter	195
123	26	26.8	27	1	VCXA_ODOSI	P49836 odontella s	196
124	26	26.8	29	2	Q9ZG21	Q9zg21 chlamydia t	197
125	26	26.8	31	2	Q7S0M7	Q7s0m7 neurospora	198
126	26	26.8	31	2	Q9XS99	Q9xsb9 ateles beiz	199
127	26	26.8	31	2	Q9KM48	Q9km48 vibrio chol	200
128	26	26.8	34	2	Q8E8Y3	Q8e8y3 shewanella	201
129	26	26.8	35	2	Q7UJMS	Q7ujm5 rhodopirell	202
130	26	26.8	35	2	Q7UUS1	Q7uus1 rhodopirell	203
131	26	26.8	35	2	Q7UJZ3	Q7uz39 rhodopirell	204
132	26	26.8	35	2	Q8EXS3	Q8exs3 leptospira	205
133	26	26.8	38	2	Q7M2T4	Q7m2t4 cryptolagus	206
134	26	26.8	38	2	Q6J5M5	Q6j5m5 streptococc	207
135	26	26.8	38	2	Q8KYR3	Q8kyr3 bacillus an	208
136	26	26.8	38	2	Q8CKU0	Q8cku0 yersinia pe	209
137	26	26.8	38	2	Q8E265	Q8e265 streptococc	210
138	26	26.8	38	2	Q8P2N0	Q8p2n0 streptococc	211
139	26	26.8	38	2	Q6EZX5	Q6eZX5 bacillus an	212
140	26	26.8	38	2	Q6VYR1	Q6vyr1 gallus gall	213
141	26	26.8	38	2	AAQ24604	Aaq24604 gallus ga	214
142	26	26.8	38	2	AAQ24606	Aaq24606 gallus ga	215
143	26	26.8	38	2	AAQ24607	Aaq24607 gallus ga	216
144	26	26.8	38	2	AAT10375	Aat10375 streptoco	217
145	26	26.8	38	2	AAT28782	Aat28782 bacillus	218
146	26	26.8	39	2	Q8N4V7	Q8n4v7 homo sapien	219
147	26	26.8	39	2	Q7VHY8	Q7vhy8 helicobacte	220
148	26	26.8	39	2	Q8FZG8	Q8fz98 brucella su	221
149	26	26.8	39	2	Q6O573	Q6o573 mus caroli	222
150	26	26.8	40	2	Q6RV10	Q6rv10 trypanosoma	223
151	26	26.8	40	2	Q87077	Q87077 simian hemo	224
152	26	26.8	40	2	Q91FA4	Q91fa4 chilo iride	225
153	26	26.8	40	2	AAE88804	Aar88804 trypanoso	226
154	26	26.8	41	2	Q7UQP8	Q7uqp8 rhodopirell	227
155	26	26.8	41	2	Q8F8H5	Q8f8h5 leptospira	228
156	26	26.8	42	2	Q33617	Q33617 streptococc	229
157	26	26.8	42	2	O50770	O50770 borrelia bu	230
158	26	26.8	42	2	Q73136	Q73136 wolbachia p	231
159	26	26.8	42	2	AA514076	Aas14076 wolbachia	232
160	26	26.8	43	2	Q43456	Q43456 homo sapien	233
161	26	26.8	43	2	Q7MEC0	Q7mec0 vibrio vuln	234
162	26	26.8	43	2	Q7UHP3	Q7uhp3 rhodopirell	235
163	26	26.8	43	2	Q9KLP2	Q9klp2 vibrio chol	236
164	26	26.8	44	2	Q9G665	Q9g665 diporiphora	237
165	26	26.8	44	2	Q8EX99	Q8ex99 leptospira	238
166	26	26.8	44	2	Q8E2K9	Q8e2k9 leptospira	239
167	26	26.8	44	2	Q8VJ61	Q8vj61 mycobacteri	240
168	26	26.8	44	2	Q85413	Q85413 rabies viru	241
169	26	26.8	44	2	Q71LZ30	Q71l30 coelognathu	242
170	26	26.8	45	2	Q7S7W6	Q7s7w6 neurospora	243
171	26	26.8	45	2	Q7PVG5	Q7pvg5 homo sapien	244
172	26	26.8	45	2	Q7YVY4	Q7yvfy4 trypanosoma	245
173	26	26.8	45	2	Q8IG46	Q8ig46 caenorhabdi	246
174	26	26.8	45	2	Q7M214	Q7m214 solanum cri	247
175	26	26.8	45	2	Q7MSAM1	Q7msam1 arabidopsis	248
176	26	26.8	45	2	Q7ULY8	Q7uly8 rhodopirell	249
177	26	26.8	45	2	Q83NY5	Q83ny5 tropheryma	250

Aan63450	2	26.8	45	2	AAAG3450	Aan63450 caenorhab
O18851	1	26.8	46	1	AOFA_SHEEP	O18851 ovis aries
Q7PJ33	2	26.8	46	2	Q7PJ33	Q7pj33 anopheles g
Q9LA74	2	26.8	46	2	Q9LA74	Q9la74 providencia
Q7W499	2	26.8	46	2	SSPN_BACSU	Q7w499 bacillus su
Q15455	2	26.8	48	1	Q15455	Q15455 homo sapien
Q85WS4	2	26.8	48	2	Q85WS4	Q85ws4 pinus korai
Q9X472	2	26.8	48	2	Q9X472	Q9x472 enterococcu
Q8Xf33	2	26.8	48	2	Q8Xf33	Q8xf33 salmonella
Q7CPA7	2	26.8	48	2	Q7CPA7	Q7cpa7 salmonella
Q8AF91	2	26.8	48	2	Q8AF91	Q8af91 mouse mamma
Q8UHS1	2	26.8	48	2	Q8UHS1	Q8uhs1 homo sapien
Q8HYK6	2	26.8	49	2	Q8HYK6	Q8hyk6 ovis aries
Q8CLI5	2	26.8	49	2	Q8CLI5	Q8cli5 yersinia pe
P88744	2	26.8	49	2	P88744	P88744 human immun
Q7RMX5	2	26.8	49	2	Q7RMX5	Q7rmx5 plasmodium
Q55072	2	26.8	50	2	Q55072	Q55072 synechocyst
Q6J191	2	26.8	50	2	Q6J191	Q6j191 neisseria m
O61191	2	26.8	50	2	O61191	O61191 lepharisma
Q8E0D2	2	26.8	50	2	Q8E0D2	Q8e0d2 streptococc
P27386	2	26.8	50	2	P27386	P27386 bacterioph
Q6EDX2	2	26.8	50	2	Q6EDX2	Q6edx2 bacterioph
Q8F8F3	2	26.8	50	2	Q8F8F3	Q8f8f3 leptospira
Q6XA02	2	26.8	50	2	Q6XA02	Q6xa02 equus cabal
Aap78692	2	26.8	50	2	Aap78692	Aap78692 equus cab
P35451	2	26.8	50	2	P35451	P35451 bos taurus
P33036	2	26.8	50	2	P33036	P33036 acinetobact
Q79CG1	2	26.8	50	2	Q79CG1	Q79cg1 bacteroides
Q9F9L1	2	26.8	50	2	Q9F9L1	Q9f9l1 pseudomonas
P18550	2	26.8	50	2	P18550	P18550 petunia hyb
Q9S444	2	26.8	50	2	Q9S444	Q9s444 pseudomonas
Q99LY6	2	26.8	50	2	Q99LY6	Q99ly6 mus musculu
Q9GJ98	2	26.8	50	2	Q9GJ98	Q9gj98 homo sapien
Q6J184	2	26.8	50	2	Q6J184	Q6j184 neisseria g
AAS16335	2	26.8	50	2	AAS16335	Aas16335 neisseria
Q9F8AL	2	26.8	50	2	Q9F8AL	Q9f8al chlamydia t
Q732G1	2	26.8	50	2	Q732G1	Q732g1 bacillus ce
AAS42856	2	26.8	50	2	AAS42856	Aas42856 bacillus
O14937	2	26.8	50	2	O14937	O14937 homo sapien
Q837M8	2	26.8	50	2	Q837M8	Q837m8 enterococcu
Q8IAC3	2	26.8	50	2	Q8IAC3	Q8iac3 bacillus ce
Q9GZQ2	2	26.8	50	2	Q9GZQ2	Q9gzq2 mus musculu
Q64854	2	26.8	50	2	Q64854	Q64854 avian adeno
Q9YPG7	2	26.8	50	2	Q9YPG7	Q9ypg7 highlands j
Q71IB6	2	26.8	50	2	Q71IB6	Q71ib6 lactobacill
O88440	2	26.8	50	2	O88440	O88440 rattus norv
Aaq07021	2	26.8	50	2	Aaq07021	Aaq07021 lactobaci
P82908	2	26.8	50	2	P82908	P82908 bos taurus
Q85NZ6	2	26.8	50	2	Q85NZ6	Q85nz6 thecabius p
Q73KJ5	2	26.8	50	2	Q73KJ5	Q73kj5 treponema d
Q87KH8	2	26.8	50	2	Q87KH8	Q87kh8 vibrio para
Q8CLQ6	2	26.8	50	2	Q8CLQ6	Q8clq6 yersinia pe
Q70225	2	26.8	50	2	Q70225	Q70225 rattus norv
Q9EC72	2	26.8	50	2	Q9EC72	Q9ec72 mus musculu
AAS12742	2	26.8	50	2	AAS12742	Aas12742 treponema
Q6GRN3	2	26.8	50	2	Q6GRN3	Q6grn3 bos taurus
Q701K0	2	26.8	50	2	Q701K0	Q701k0 spiroplasma
Q701K1	2	26.8	50	2	Q701K1	Q701k1 spiroplasma
Q8E9Z1	2	26.8	50	2	Q8E9Z1	Q8e9z1 shewanella
AAS20601	2	26.8	50	2	AAS20601	Aas20601 bos tauru
CAF32403	2	26.8	50	2	CAF32403	CAF32403 spiroplas
Q07588	2	26.8	50	2	Q07588	Q07588 bacillus su
Q6H209	2	26.8	50	2	Q6H209	Q6h209 bacillus an
Q8EYU7	2	26.8	50	2	Q8EYU7	Q8eyu7 leptospira
Q8FBU6	2	26.8	50	2	Q8FBU6	Q8fbu6 leptospira
Q9PC23	2	26.8	50	2	Q9PC23	Q9pc23 xyella fas
Q8ULK4	2	26.8	50	2	Q8ULK4	Q8ulk4 pyrococcus
Q7RBN2	2	26.8	50	2	Q7RBN2	Q7rbn2 plasmodium
Q95Z12	2	26.8	50	2	Q95Z12	Q95z12 ceratitis r
Q55239	2	26.8	50	2	Q55239	Q55239 synechocyst
Q8F1K6	2	26.8	50	2	Q8F1K6	Q8f1k6 leptospira

251 25 25.8 37 2 Q6EAV7
 252 25 25.8 37 2 Q6EAW0
 253 25 25.8 37 2 Q6EAW3
 254 25 25.8 38 2 Q7RK19
 255 25 25.8 38 2 Q8MTP9
 256 25 25.8 38 2 Q9G051
 257 25 25.8 38 2 Q7MVQ3
 258 25 25.8 38 2 Q8CM05
 259 25 25.8 38 2 Q8CMJ4
 260 25 25.8 38 2 Q8CMJ5
 261 25 25.8 38 2 Q8EBE9
 262 25 25.8 39 2 Q8MSJ8
 263 25 25.8 39 2 Q7Z2U6
 264 25 25.8 39 2 Q738P5
 265 25 25.8 39 2 AAS41267
 266 25 25.8 39 2 AAS43472
 267 25 25.8 40 1 MSDS_DROME
 268 25 25.8 40 2 Q7RQH7
 269 25 25.8 40 2 Q6FD11
 270 25 25.8 40 2 Q6EAV4
 271 25 25.8 40 2 Q6EAW9
 272 25 25.8 41 2 Q26188
 273 25 25.8 41 2 Q8IUB2
 274 25 25.8 41 2 Q8CLB1
 275 25 25.8 41 2 Q6EAV4
 276 25 25.8 42 1 POAL_SALKA
 277 25 25.8 42 1 SPI8_SOUTU
 278 25 25.8 42 2 Q7PHZ3
 279 25 25.8 42 2 Q8CLQ1
 280 25 25.8 43 2 Q7PF17
 281 25 25.8 43 2 Q8L1G0
 282 25 25.8 43 2 Q8E275
 283 25 25.8 43 2 Q91J96
 284 25 25.8 44 1 SSPN_BACAN
 285 25 25.8 44 1 SSPN_BACC1
 286 25 25.8 44 1 SSPN_BACCR
 287 25 25.8 44 2 Q85X60
 288 25 25.8 44 2 Q9SLP0
 289 25 25.8 44 2 Q8KYL0
 290 25 25.8 44 2 Q9X160
 291 25 25.8 44 2 Q6E2M6
 292 25 25.8 44 2 Q04061
 293 25 25.8 44 2 Q92318
 294 25 25.8 44 2 AAS42534
 295 25 25.8 44 2 AAT28893
 296 25 25.8 44 2 AAT32779
 297 25 25.8 45 2 Q7REM6
 298 25 25.8 45 2 Q49768
 299 25 25.8 45 2 Q7DLH4
 300 25 25.8 45 2 Q50969
 301 25 25.8 45 2 Q81QU3
 302 25 25.8 45 2 Q7ZFT5
 303 25 25.8 45 2 AAT31445
 304 25 25.8 46 1 YC12_SYNEL
 305 25 25.8 46 1 YCX2_ODOS1
 306 25 25.8 46 2 O15745
 307 25 25.8 46 2 O49699
 308 25 25.8 46 2 Q6TFF1
 309 25 25.8 46 2 Q7VNF6
 310 25 25.8 46 2 Q8C9V2
 311 25 25.8 46 2 Q91F83
 312 25 25.8 46 2 AAR87108
 313 25 25.8 47 1 VG60_BPML5
 314 25 25.8 47 2 Q9TWY9
 315 25 25.8 47 2 Q8E116
 316 25 25.8 47 2 Q8F5F0
 317 25 25.8 48 2 Q75253
 318 25 25.8 48 1 LECA_DLOOR
 319 25 25.8 49 1 RL40_METAC
 320 25 25.8 49 2 Q9UD55
 321 25 25.8 49 2 Q8MR92
 322 25 25.8 49 2 Q8MR92
 323 25 25.8 49 2 Q8MR92

324 25 25.8 49 2 Q6LAE3
 325 25 25.8 49 2 Q59529
 326 25 25.8 49 2 Q81N12
 327 25 25.8 49 2 Q91J48
 328 25 25.8 49 2 AAT78647
 329 25 25.8 49 2 AAT32515
 330 25 25.8 50 2 Q6BS22
 331 25 25.8 50 2 Q7Y3M7
 332 25 25.8 50 2 Q81SF8
 333 25 25.8 50 2 Q88FK4
 334 25 25.8 50 2 Q92HH9
 335 25 25.8 50 2 Q9QX44
 336 25 25.8 50 2 Q9QX44
 337 25 25.8 50 2 Q9R074
 338 25 25.8 50 2 Q9R077
 339 25 25.8 50 2 Q8ORE1
 340 25 25.8 50 2 Q81J44
 341 25 25.8 50 2 Q91JA6
 342 25 25.8 50 2 Q8V9K7
 343 25 25.8 50 2 Q8V9L0
 344 25 25.8 50 2 AAT30811
 345 24.5 25.3 50 2 Q7PCT2
 346 24.5 25.3 28 1 ETX2_BACCE
 347 24.5 25.3 31 2 Q6C7X1
 348 24.5 25.3 31 2 Q73HP3
 349 24.5 25.3 31 2 AAS14220
 350 24.5 25.3 32 2 Q9NGM5
 351 24.5 25.3 32 2 Q97RN0
 352 24.5 25.3 38 2 Q36273
 353 24.5 25.3 38 2 Q79SD6
 354 24.5 25.3 42 2 Q8VNP2
 355 24.5 25.3 42 2 Q72314
 356 24.5 25.3 42 2 AAS96839
 357 24.5 25.3 46 2 Q7X4K2
 358 24.5 25.3 47 2 Q81DJ1
 359 24.5 25.3 50 1 S10D_CHICK
 360 24.5 25.3 50 2 Q9UJ26
 361 24 24.7 19 2 Q9BMB9
 362 24 24.7 19 2 Q7GEF1
 363 24 24.7 19 2 Q7GEF2
 364 24 24.7 19 2 Q36925
 365 24 24.7 19 2 Q9G466
 366 24 24.7 20 2 Q9TWN3
 367 24 24.7 20 2 Q9TWR0
 368 24 24.7 20 2 Q36584
 369 24 24.7 23 2 Q36806
 370 24 24.7 23 2 Q63334
 371 24 24.7 23 2 Q04280
 372 24 24.7 25 1 QCE1_LEPOE
 373 24 24.7 25 2 Q9BGJ1
 374 24 24.7 25 2 Q714T3
 375 24 24.7 25 2 Q60839
 376 24 24.7 25 2 Q91455
 377 24 24.7 25 2 AAO12673
 378 24 24.7 26 2 Q75BX0
 379 24 24.7 26 2 Q6R9Y6
 380 24 24.7 26 2 Q6R9Y7
 381 24 24.7 26 2 Q6R9Y8
 382 24 24.7 26 2 Q6R9Z1
 383 24 24.7 26 2 Q9R884
 384 24 24.7 26 2 AAS88092
 385 24 24.7 26 2 AAS88095
 386 24 24.7 26 2 AAS88096
 387 24 24.7 26 2 AAS88097
 388 24 24.7 27 2 Q7RYX5
 389 24 24.7 27 2 Q85XP5
 390 24 24.7 28 2 Q71PB1
 391 24 24.7 28 2 Q36593
 392 24 24.7 28 2 AAO03619
 393 24 24.7 29 1 GLUC_CHIBR
 394 24 24.7 29 2 Q13782
 395 24 24.7 29 2 Q8HZ61
 396 24 24.7 29 2 Q92126
 397 24 24.7 30 2 Q36780

Q61ae3 neisseria m
 Q59529 mycobacteri
 Q81n12 bacillus an
 Q91j48 hepatitis c
 Aat78647 neisseria
 Aat32515 bacillus
 Q6bs22 debaryomyce
 Q7y3m7 enterobacte
 Q81sf8 bacillus an
 Q88fk4 pseudomonas
 Q92hb9 rickettsia
 Q9qx44 cricetidae
 Q9r074 mus musculu
 Q9r077 cavia porce
 Q80rel tt virus. o
 Q1j44 hepatitis c
 Q91ja6 hepatitis c
 Q8v9k7 human papil
 Q8v9l0 human papil
 Aat30811 bacillus
 Q7pct2 plasmodium
 P80568 bacillus ce
 Q6c7x1 varrowia li
 Q73hp3 wolbachia p
 Aas14220 wolbachia
 Q9ngm5 strongyloce
 Q36273 streptococc
 Q79sd6 streptococ
 Q8vnp2 escherichia
 Q72314 desulfovibr
 Aas96839 desulfovi
 Q7x4k2 streptococc
 Q81dj1 bacillus ce
 P51964 gallus gall
 Q9u26 schizosacch
 Q9bmb9 ceratitits c
 Q7gef1 nicotiana v
 Q7gef2 nicotiana a
 Q36925 nicotiana p
 Q9g466 brassica ju
 Q9twn3 thelleria s
 Q9twr0 blattella g
 Q36584 nicotiana g
 Q36806 solanum tub
 Q63334 rattus norv
 Q04280 african swi
 P83951 leptodactyl
 Q9bgj1 tarsius ban
 Q714t3 skeletonema
 Q60839 mus musculu
 Q91455 sericornis
 Aaq12673 skeleton
 Q7rbx0 plasmodium
 Q6ry6 trachypithe
 Q6ry7 trachypithe
 Q6ry8 pygathrix n
 Q6ry21 pygathrix b
 Q9r884 chlamydia t
 Aas88092 pygathrix
 Aas88095 pygathrix
 Aas88096 trachypit
 Aas88097 trachypit
 Q7ryx5 neuropeora
 Q85xp5 drosophila
 Q71pb1 aerobryidiu
 Q36593 nicotiana p
 Aaq03619 aerobryid
 P31297 chinchilla
 Q13782 homo sapien
 Q8hz61 bos taurus
 Q92126 mus musculu
 Q36780 lycopersico

397	24	24.7	30	2	Q8EG77	Q8eg77 shewanella
398	24	24.7	30	2	Q8EGQ3	Q8egq3 shewanella
399	24	24.7	31	2	Q34922	Q34922 limulus pol
400	24	24.7	31	2	Q6SFB9	Q6sfb9 uncultured
401	24	24.7	31	2	Q8RTS5	Q8rts5 uncultured
402	24	24.7	31	2	Q97NU1	Q97nu1 streptococc
403	24	24.7	31	2	Q8JU93	Q8ju93 ramphocincl
404	24	24.7	31	2	Q8JU96	Q8ju96 minus gilvu
405	24	24.7	31	2	Q8JU97	Q8ju97 melanotus c
406	24	24.7	31	2	Q8JU98	Q8ju98 melanoptila
407	24	24.7	31	2	Q8JU99	Q8ju99 margarops f
408	24	24.7	31	2	Q8JU9A	Q8ju9a margarops f
409	24	24.7	31	2	Q8JU9B	Q8ju9b margarops f
410	24	24.7	31	2	Q8JU9C	Q8ju9c margarops f
411	24	24.7	31	2	Q8JU9D	Q8ju9d margarops f
412	24	24.7	31	2	Q8JU9E	Q8ju9e margarops f
413	24	24.7	31	2	Q8JU9F	Q8ju9f margarops f
414	24	24.7	31	2	Q8JU9G	Q8ju9g margarops f
415	24	24.7	31	2	Q8JU9H	Q8ju9h margarops f
416	24	24.7	31	2	Q8JU9I	Q8ju9i margarops f
417	24	24.7	31	2	Q8JU9J	Q8ju9j margarops f
418	24	24.7	31	2	Q8JU9K	Q8ju9k margarops f
419	24	24.7	31	2	Q8JU9L	Q8ju9l margarops f
420	24	24.7	31	2	Q8JU9M	Q8ju9m margarops f
421	24	24.7	31	2	Q8JU9N	Q8ju9n margarops f
422	24	24.7	31	2	Q8JU9O	Q8ju9o margarops f
423	24	24.7	31	2	Q8JU9P	Q8ju9p margarops f
424	24	24.7	31	2	Q8JU9Q	Q8ju9q margarops f
425	24	24.7	31	2	Q8JU9R	Q8ju9r margarops f
426	24	24.7	31	2	Q8JU9S	Q8ju9s margarops f
427	24	24.7	31	2	Q8JU9T	Q8ju9t margarops f
428	24	24.7	31	2	Q8JU9U	Q8ju9u margarops f
429	24	24.7	31	2	Q8JU9V	Q8ju9v margarops f
430	24	24.7	31	2	Q8JU9W	Q8ju9w margarops f
431	24	24.7	31	2	Q8JU9X	Q8ju9x margarops f
432	24	24.7	31	2	Q8JU9Y	Q8ju9y margarops f
433	24	24.7	31	2	Q8JU9Z	Q8ju9z margarops f
434	24	24.7	31	2	Q8JU9A	Q8ju9a margarops f
435	24	24.7	31	2	Q8JU9B	Q8ju9b margarops f
436	24	24.7	31	2	Q8JU9C	Q8ju9c margarops f
437	24	24.7	31	2	Q8JU9D	Q8ju9d margarops f
438	24	24.7	31	2	Q8JU9E	Q8ju9e margarops f
439	24	24.7	31	2	Q8JU9F	Q8ju9f margarops f
440	24	24.7	31	2	Q8JU9G	Q8ju9g margarops f
441	24	24.7	31	2	Q8JU9H	Q8ju9h margarops f
442	24	24.7	31	2	Q8JU9I	Q8ju9i margarops f
443	24	24.7	31	2	Q8JU9J	Q8ju9j margarops f
444	24	24.7	31	2	Q8JU9K	Q8ju9k margarops f
445	24	24.7	31	2	Q8JU9L	Q8ju9l margarops f
446	24	24.7	31	2	Q8JU9M	Q8ju9m margarops f
447	24	24.7	31	2	Q8JU9N	Q8ju9n margarops f
448	24	24.7	31	2	Q8JU9O	Q8ju9o margarops f
449	24	24.7	31	2	Q8JU9P	Q8ju9p margarops f
450	24	24.7	31	2	Q8JU9Q	Q8ju9q margarops f
451	24	24.7	31	2	Q8JU9R	Q8ju9r margarops f
452	24	24.7	31	2	Q8JU9S	Q8ju9s margarops f
453	24	24.7	31	2	Q8JU9T	Q8ju9t margarops f
454	24	24.7	31	2	Q8JU9U	Q8ju9u margarops f
455	24	24.7	31	2	Q8JU9V	Q8ju9v margarops f
456	24	24.7	31	2	Q8JU9W	Q8ju9w margarops f
457	24	24.7	31	2	Q8JU9X	Q8ju9x margarops f
458	24	24.7	31	2	Q8JU9Y	Q8ju9y margarops f
459	24	24.7	31	2	Q8JU9Z	Q8ju9z margarops f
460	24	24.7	31	2	Q8JU9A	Q8ju9a margarops f
461	24	24.7	31	2	Q8JU9B	Q8ju9b margarops f
462	24	24.7	31	2	Q8JU9C	Q8ju9c margarops f
463	24	24.7	31	2	Q8JU9D	Q8ju9d margarops f
464	24	24.7	31	2	Q8JU9E	Q8ju9e margarops f
465	24	24.7	31	2	Q8JU9F	Q8ju9f margarops f
466	24	24.7	31	2	Q8JU9G	Q8ju9g margarops f
467	24	24.7	31	2	Q8JU9H	Q8ju9h margarops f
468	24	24.7	31	2	Q8JU9I	Q8ju9i margarops f
469	24	24.7	31	2	Q8JU9J	Q8ju9j margarops f

ALIGNMENTS

RESULT 1

Q7R8B1	PRELIMINARY;	PRT;	50 AA.
AC	Q7R8B1		
DT	01-MAR-2004 (TRENBLrel. 26, Created)		
DT	01-MAR-2004 (TRENBLrel. 26, Last sequence update)		
DT	01-MAR-2004 (TRENBLrel. 26, Last annotation update)		
DE	Hypothetical protein (fragment).		
GN	Name=PY07312.		
OS	Plasmodium yoelii yoelii.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=73239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=17XNL;		
RX	PubMed=1236865;		
RA	Carlton J.M., Angiuoli S.V., Suh B.B., Koolij T.W., Perteza M., Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L., Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.D., Shallom S.J., van Alen S.E., Riedmuller S.B., Feldblyum T.V., Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M., Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A., Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B., van Lin L.H., Jause C.J., Waters A.P., Smith H.O., White O.R., Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J., Carucci D.J.;		
RT	"Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii.";		
RL	Nature 419:512-519(2002).		
CC	-!- CAUTION: the sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.		
CC	EMBL:AA01002650; EAA19712.1; -.		
DR	Hypothetical protein.		
KW	Hypothetical protein.		
FT	NON_TER		
SQ	SEQUENCE 50 AA; 5782 MW; 69852D6DB5C33622 CRC64;		


```

Query Match      39.2%; Score 38; DB 2; Length 50;
Best Local Similarity 46.2%; Pred. No. 1.1e+02;
Matches      6; Conservative      1; Mismatches      6; Indels      0; Gaps      0;

QY      1 EPNHNSKIAPKI 13
      ||||| :| :| :|
Db      31 EPNHFNKNTMMKV 43

RESULT 2
Q6JDL7 PRELIMINARY; PRT; 20 AA.
AC Q6JDL7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Desmoglein 3 preprotein (Fragment).
GN Name=DSG3;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15233990;
RA Housley D.J.E., Ritzert E., Venta P.J.;
RT "Comparative radiation hybrid map of canine chromosome 1 incorporating
RL SNPs and indel polymorphisms.";
RL Genomics 84:248-264(2004).
DR EMBL; AY514706; AAT44553.1; -.
FT NON_TER 1
FT NON_TER 20
SQ SEQUENCE 20 AA; 2168 MW; 6A704F9859DD68D6 CRC64;

Query Match      37.1%; Score 36; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 93;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 EPNHLN 6
      |||||
Db      15 EPNHLN 20

RESULT 3
Q8FFB7 PRELIMINARY; PRT; 41 AA.
AC Q8FFB7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein c2949.
GN OrderedLocusNames=c2949;
OS Escherichia coli O6
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22382234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.J., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AB016764; AAN81399.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 41 AA; 4598 MW; 15F1C1B595551F33 CRC64;

Query Match      36.6%; Score 35.5; DB 2; Length 41;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;

QY      2 PNHLNS--KIAPKIVSQEP 18
      | : ||| | :| :| :|
Db      1 PHLNHLKLGFDVIAEP 19

RESULT 5
CAB36907 PRELIMINARY; PRT; 21 AA.
AC CAB36907
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Caveolin-2 (Fragment).
GN CAVEOLIN-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fra A.M., Mastroianni N., Mancini R., Pasqualetto E., Sitia R.;
RT "Human caveolin-1 and caveolin-2 are closely-linked genes colocalised
RT with WI-5336 in a region of 7q31 frequently deleted in tumors.";
RL Genomics 56:355-356(1999).
DR EMBL; AJ011300; CAB36907.1; -.
DR InterPro; IPR001612; Caveolin.
DR PROSITE; PS01210; CAVEOLIN; 1.
FT NON_TER 1
FT NON_TER 21
SQ SEQUENCE 21 AA; 2373 MW; 4AC6FC4BB5C265E4 CRC64;

Query Match      36.1%; Score 35; DB 2; Length 21;
Best Local Similarity 47.4%; Pred. No. 1.4e+02;
Matches      9; Conservative      3; Mismatches      5; Indels      2; Gaps      1;

QY      2 PNHLNS--KIAPKIVSQEP 18
      | : ||| | :| :| :|
Db      1 PHLNHLKLGFDVIAEP 19

RESULT 5
CAB36907 PRELIMINARY; PRT; 21 AA.
AC CAB36907
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Caveolin-2 (Fragment).
GN CAVEOLIN-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fra A.M., Mastroianni N., Mancini R., Pasqualetto E., Sitia R.;
RT "Human caveolin-1 and caveolin-2 are closely-linked genes colocalised
RT with WI-5336 in a region of 7q31 frequently deleted in tumors.";
RL Genomics 56:355-356(1999).
DR EMBL; AJ011300; CAB36907.1; -.
FT NON_TER 1
FT NON_TER 21
SQ SEQUENCE 21 AA; 2373 MW; 4AC6FC4BB5C265E4 CRC64;

Query Match      36.1%; Score 35; DB 2; Length 21;
Best Local Similarity 47.4%; Pred. No. 1.4e+02;
Matches      9; Conservative      3; Mismatches      5; Indels      2; Gaps      1;

QY      2 PNHLNS--KIAPKIVSQEP 18
      | : ||| | :| :| :|

```

1 PHRLNSHLKLGFEVDIAEP 19

```

RESULT 6
O85326 PRELIMINARY; PRT; 38 AA.
ID
AC O85326;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein sp14_Q.
GN Name=sp14_Q;
OS Salmonella Typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=98298059; Pubmed=9632606;
RA Wong K.K., McClelland M., Stillwell L.C., Sisk E.C., Thurston S.J.,
RA Saffer J.D.;
RT "Identification and sequence analysis of a 27-kilobase chromosomal
RT fragment containing a Salmonella pathogenicity island located at 92
RT minutes on the chromosome map of Salmonella enterica serovar
RT typhimurium LT2."
RL Infect. Immun. 66:3365-3371(1998) .
DR EMBL; AF060869; AAC26654.1; -.
DR F01; F14885;
DR Hypothetical protein.
KW SEQUENCE 38 AA; 4452 MW; C977BF84F9C51846 CRC64;
SQ

```

```

Query Match      35.1%; Score 34; DB 2; Length 38;
Best Local Similarity 46.7%; Pred.No 4e+02;
Matches 7; Conservative 4; Mismatches 0; Gaps 0;
QY 1 BPNNHNSKIAFKIVS 15
DB 6 EPYLSAETALSIVS 20

```

```

RESULT 7
Q9PSI5 PRELIMINARY; PRT; 20 AA.
ID Q9PSI5
AC Q9PSI5;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DT 01-MAY-2003 (TRENBLrel. 24, Last annotation update)
DE Creatine kinase (EC 2.7.3.2) (Fragment).
OS Oncorhynchus kisutch (Coho salmon)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8019;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93080727; PubMed=1449598;
RA White K.C., Rabbitt P.C., Buechter D.D., Kenyon G.L.;
RT "the principal islet of the Coho salmon (Oncorhynchus kisutch) contains
RT the BB isoenzyme of creatine kinase.";
RL J. Protein Chem. 11:489-494(1992).
RL PIR; A53875; A53875.
SQ GO; GO:0004111; F:creatine kinase activity; IEA.
SO SEQUENCE 20 AA; 2425 MW; BFBCT71D62DA86E11 CRC64;

```

Query Match	34.0%;	Score 33;	DB 2;	Length 20;
Best Local Similarity	40.0%;	Pred. No. 3e+02;		
Matches	6;	Conservative 3;	Mismatches 6;	Indels 0;
Gaps				0;

```

RESULT 8
C8FEPS
ID Q8FEPS PRELIMINARY; PRT; 41 AA.
AC Q8FEPS;
DT 01-WAR-2003 (TremBLrel. 23, Created)
DT 01-WAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-WAR-2003 (TremBLrel. 23, Last annotation update)
DE Hypothetical protein c3248.
GN OrderedLocusNames=c3248;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:HI / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Licu S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002) .
RW EMBL; AF016765; AAN01700.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 41 AA; 5081 MW; 346DA306456E56C9 CRC64;

Query Match 34.0%; Score 33; DB 2; Length 41;
Best Local Similarity 46.2%; Pred. No. 6.4e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps:

```

RESULT 9	PRELIMINARY:	PRT:	50 AA.
Q9PCZ2			
ID	Q9PCZ2		
AC	Q9PCZ2;		
DT	01-OCT-2000 (trEMBLrel. 15, Created)		
DT	01-OCT-2000 (trEMBLrel. 15, Last sequence update)		
DT	01-JUN-2003 (trEMBLrel. 24, Last annotation update)		
DE	Hypothetical protein.		
GN	OrderedLocusNames=Xfi1612;		
OS	<i>Xyella fastidiosa</i> .		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;		
OC	Xanthomonadaceae; <i>Xyella</i> .		
OX	NCBI_TaxID=2371;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=98ac;		
RX	MEDLINE=20365717; PubMed=10910347;		
RA	Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,		
RA	Alvarenga R., Alves L.M.C., Araya J.B., Baia G.S., Baptista C.S.,		
RA	Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,		
RA	Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,		
RA	Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,		
RA	Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,		
RA	Facincini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,		
RA	Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,		
RA	Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,		
RA	Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,		
RA	Krieger J.E., Kuramae E.E., Laigret P., Lambais M.R., Leite L.C.C.,		
RA	Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,		
RA	Machado M.A., Madeira A.M.B., Madeira H.M.F., Marino C.L.,		
RA	Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,		
RA	Menck C.F.M., Miracca E.C., Miyaki C.I., Monteiro-Vitorello C.B.,		
RA	Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,		
RA	Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,		
RA	de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,		

```

RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawaaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE003988; AAF84421.1; -.
DR PIR; D82660; D82660.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 50 AA; 5754 MW; 57A05A2E5E6F956C CRC64;

Query Match 34.0%; Score 33; DB 2; Length 50;
Best Local Similarity 46.7%; Pred. No. 7.9e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSQEPA 19
||| : : : :
Db 23 LNKKLSNKIKNEAA 37

RESULT 10
Q9RHG5 PRELIMINARY; PRT; 35 AA.
AC Q9RHG5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Organomercurial lyase (Fragment).
GN Name=merB3;
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89123092; PubMed=2536669;
RA Wang Y., Moore M., Levinson H.S., Silver S., Walsh C., Mahler I.;
RT "Nucleotide sequence of a chromosomal mercury resistance determinant
from a Bacillus sp. with broad-spectrum mercury resistance.";
RL J. Bacteriol. 171:83-92(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Huang C., Narita M., Ishii H., Suzuki T., Endo G.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036431; BAA89374.1; -.
DR GO; GO:0016829; F-lyase activity; IEA.
KW Lyase.
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3894 MW; 806A31F5F34DF864 CRC64;

Query Match 33.0%; Score 32; DB 2; Length 35;
Best Local Similarity 41.7%; Pred. No. 8e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 NHLNSKIAPKIV 14
|||| : : :
Db 4 NHLNLSLKDKVL 15

RESULT 11
Q7LYB7 PRELIMINARY; PRT; 37 AA.
AC Q7LYB7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE VhtC protein (Fragment).
GN Name=vhtC;
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;

```

```

OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goi;
RX MEDLINE=96106133; PubMed=8572889;
RA Deppenmeier U.;
RT "Different structure and expression of the operons encoding the
membrane-bound hydrogenases from Methanosarcina mazei G01.";
RL Arch. Microbiol. 164:370-376(1995).
DR EMBL; X91851; CAA62961.1; -.
FT NON_TER 1 1
SQ SEQUENCE 37 AA; 4428 MW; DCD7CA091270FF49 CRC64;

Query Match 33.0%; Score 32; DB 2; Length 37;
Best Local Similarity 33.3%; Pred. No. 8.5e+02;
Matches 6; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 PNHLNSKIAPKIVSQEPA 19
|||| : : : :
Db 16 PNHLPDRELWDFSDKPS 33

RESULT 12
Q84XN8 PRELIMINARY; PRT; 41 AA.
ID Q84XN8
AC Q84XN8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dehydration responsive protein (Fragment).
OS Gossypium barbadense (Sea-island cotton) (Egyptian cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Gossypium.
OX NCBI_TaxID=3634;
RN [1]
RP SEQUENCE FROM N.A.
RA Adams K.L., Percifield R.J., Wendel J.F.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY185523; AAO33906.1; -.
DR InterPro; IPR004873; BURP.
DR Pfam; PF03181; BURP; 1.
FT NON_TER 1 1
SQ SEQUENCE 41 AA; 4704 MW; 735501EC8C090087 CRC64;

Query Match 33.0%; Score 32; DB 2; Length 41;
Best Local Similarity 41.2%; Pred. No. 9.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 2 PNHLNSKIAPKIVSQEP 18
|||| : : : :
Db 10 PKHL----AFQVLKVEP 22

RESULT 13
Q84XN9 PRELIMINARY; PRT; 41 AA.
ID Q84XN9
AC Q84XN9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dehydration responsive protein (Fragment).
OS Gossypium barbadense (Sea-island cotton) (Egyptian cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Gossypium.
OX NCBI_TaxID=3634;
RN [1]
RP SEQUENCE FROM N.A.
RA Adams K.L., Percifield R.J., Wendel J.F.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY185522; AAO33905.1; -.

```

DR	InterPro; IPR004873; BURP.	
DR	Pfam; PF03181; BURP; 1.	
FT	NON TER	1
SQ	SEQUENCE	41 AA; 4719 MW; 6539D1EC8C090087 CRC64;
Query Match		
Best Local Similarity		
Matches		
7; Conservative		
3; Mismatches		
3; Indels		
4; Gaps		
1;		
RESULT 14		
Qy	2 PNHLNSKIAFKIVSQEP 18	
Db	10 PKHL-----AFQVLKVEP 22	
SEQUENCE		
41 AA; 4719 MW; 6539D1EC8C090087 CRC64;		
Query Match		
Best Local Similarity		
Matches		
7; Conservative		
3; Mismatches		
3; Indels		
4; Gaps		
1;		
RESULT 15		
Qy	2 PNHLNSKIAFKIVSQEP 18	
Db	10 PKHL-----AFQVLKVEP 22	
SEQUENCE		
41 AA; 4719 MW; 6539D1EC8C090087 CRC64;		
Query Match		
Best Local Similarity		
Matches		
7; Conservative		
3; Mismatches		
3; Indels		
4; Gaps		
1;		
RESULT 16		
Qy	2 PNHLNSKIAFKIVSQEP 18	
Db	10 PKHL-----AFQVLKVEP 22	
SEQUENCE		
41 AA; 4719 MW; 6539D1EC8C090087 CRC64;		
Query Match		
Best Local Similarity		
Matches		
7; Conservative		
3; Mismatches		
3; Indels		
4; Gaps		
1;		
RESULT 17		
Qy	3 NNLNSKIAFKIVS 15	
Db	2 NSLSILHYKIIA 14	
SEQUENCE		
41 AA; 4816 MW; 8C6A4D7814E1598B CRC64;		
Query Match		
Best Local Similarity		
Matches		
6; Conservative		
4; Mismatches		
3; Indels		
0; Gaps		
0;		
RESULT 18		
Qy	2 PNHLNSKIAFKIVSQEP 18	
Db	10 PKHL-----AFQVLKVEP 22	
SEQUENCE		
41 AA; 4719 MW; 6539D1EC8C090087 CRC64;		
Query Match		
Best Local Similarity		
Matches		
7; Conservative		
3; Mismatches		
3; Indels		
4; Gaps		
1;		
RESULT 19		
Qy	6 NSKIAFKIVSQEP 19	
SEQUENCE		
31 AA; 3658 MW; A7E9C61F5DC6F6CE6 CRC64;		
Query Match		
Best Local Similarity		
Matches		
5; Conservative		
5; Mismatches		
3; Indels		
0; Gaps		
0;		

```

Db      |:::| |:::|
9 NKEVSVKWTQDP 21

RESULT 18
Q8E187 PRELIMINARY; PRT; 31 AA.
AC Q8E187
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SO0957.
GN OrderedLocuNames=SO0957;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbr749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Meehe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.F., Umayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RL Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR ENBL; AE015541; AAN54031.1; -
DR TIGR; SO0957; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 31 AA; 3377 MW; BC11F46F8B60571 CRC64;

Query Match 32.0%; Score 31; DB 2; Length 31;
Best Local Similarity 62.5%; Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLNSKI 9
Db 9 PNHLATKL 16

RESULT 19
Q7UEV7 PRELIMINARY; PRT; 43 AA.
AC Q7UEV7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=RBI0531;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR ENBL; BX294151; CAD79927.1; -
DR InterPro; IPR011479; DUF1586.
DR Pfam; PF07625; DUF1586; 1.

KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 43 AA; 4746 MW; 312A586866FCE3D7 CRC64;

Query Match 32.0%; Score 31; DB 2; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSOEP 18
Db 9 LTRTAPAAVKOMP 22

RESULT 20
Q8CKB8 PRELIMINARY; PRT; 50 AA.
AC Q8CKB8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical
GN OrderedLocuNames=y3768;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Petherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
DR ENBL; AE013980; AAM87313.1; -
KW Hypothetical protein.
SQ SEQUENCE 50 AA; 6023 MW; D54979A6AD1132FD CRC64;

Query Match 32.0%; Score 31; DB 2; Length 50;
Best Local Similarity 53.8%; Pred. No. 1.7e+03;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 NHLNSKIAPKIVS 15
Db 8 NLNKKILIKICS 20

RESULT 21
RS7_STAAU STANDARD; PRT; 42 AA.
AC P48940;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 30S ribosomal protein S7 (fragment).
GN Name=rpS6;
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTC 8325;
RX MEDLINE=98241542; PubMed=9573165;
RA Wada A., Watanabe H.;
RT "penicillin-binding protein 1 of Staphylococcus aureus is essential
RT for growth."
RL J. Bacteriol. 180:2759-2765(1998).
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the head
CC domain of the 30S subunit. Is located at the subunit interface
CC close to the decoding center, probably blocks exit of the E-site

```

```

CC      tRNA (By similarity).
CC      -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9
CC      and S11 (By similarity).
CC      -1- SIMILARITY: Belongs to the S7P family of ribosomal proteins.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U20869; AAC4354.1; -.
CC      DR      HSSP; P22744; IHUS.
CC      DR      HAMAP; MF 00480; -.
CC      DR      InterPro; IPR000235; Ribosomal S7.
CC      DR      Pfam; PF00177; Ribosomal S7; 1.
CC      DR      PROSITE; PS00052; RIBOSOMAL S7; PARTIAL.
CC      KW      Ribosomal protein; RNA-binding; rRNA-binding; tRNA-binding.
CC      FT      NON_TER 42
CC      SQ      SEQUENCE 42 AA; 4723 MW; 06785303ED186145 CRC64;
CC
CC      Query Match 31.4%; Score 30.5; DB 1; Length 42;
CC      Best Local Similarity 37.5%; Pred. No. 1.8e+03;
CC      Matches 6; Conservative 6; Mismatches 3; Indels 1; Gaps 1;
CC
CC      QY      1 EPNHLSKIAFKIVSQ 16
CC      Db      15 DPIH-NSKLVTKLINK 29
CC
CC      RESULT 22
CC      Q9W687 PRELIMINARY; PRT; 45 AA.
CC      AC      Q9W687;
CC      DT      01-NOV-1999 (TrEMBLrel. 12, Created)
CC      DT      01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
CC      DT      01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
CC      DE      Proenkephalin (Fragment).
CC      OS      Acipenser transmontanus (White sturgeon).
CC      OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      AC      Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
CC      AC      Acipenser.
CC      OX      NCBI_TaxID=7904;
CC      RN      [1]
CC      RP      SEQUENCE FROM N.A.
CC      RC      TISSUE=Brain;
CC      RA      Danielson P.B., Dore R.M.;
CC      RL      Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
CC      DR      EMBL; AF084454; AAD29129.1; -.
CC      FT      NON_TER 1
CC      SQ      SEQUENCE 45 AA; 5393 MW; 86F50BBA2FBE646F CRC64;
CC
CC      Query Match 31.4%; Score 30.5; DB 2; Length 45;
CC      Best Local Similarity 41.2%; Pred. No. 1.9e+03;
CC      Matches 7; Conservative 4; Mismatches 3; Indels 3; Gaps 1;
CC
CC      QY      1 EPNHLSKIAFKIVSQE 17
CC      Db      21 QPEHTDSK---EITSEE 34
CC
CC      RESULT 23
CC      Q7RM25 PRELIMINARY; PRT; 31 AA.
CC      AC      Q7RM25;
CC      DT      01-MAR-2004 (TrEMBLrel. 26, Created)
CC      DT      01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
CC      DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
CC      DE      Hypothetical protein (Fragment).
CC      GN      Name=PY02363;
CC      OS      Plasmodium yoelii yoelii.

```

```

OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=73239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=17XNL;
RX      PubMed=12368865;
RA      Carlton J.M., Angiuoli S.V., Suh B.B., Koolij T.W., Perteau M.,
RA      Silva J.C., Ermolaeva M.D., Allen J.E., Sengenut J.D., Koo H.L.,
RA      Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA      Shalom S.J., van Aken S.B., Riedmuller S.B., Feldblyum T.V.,
RA      Chao J.K., Quackenbush J., Sedegah M., Shoabai A., Cummings L.M.,
RA      Flores L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA      Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA      van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA      Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA      Carucci D.J.;
RA      "Genome sequence and comparative analysis of the model rodent malaria
RA      parasite Plasmodium yoelii yoelii."
RL      Nature 419:512-519(2002).
CC      -1- CAUTION: The sequence shown here is derived from an
CC      EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC      preliminary data.
CC      DR      EMBL; AABL01000645; EAA21801.1; -.
CC      KW      Hypothetical protein.
CC      FT      NON_TER 1
CC      SQ      SEQUENCE 31 AA; 3601 MW; 143C117535E03512 CRC64;
CC
CC      Query Match 30.9%; Score 30; DB 2; Length 31;
CC      Best Local Similarity 38.5%; Pred. No. 1.6e+03;
CC      Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
CC
CC      QY      1 EPNHLSKIAFKI 13
CC      Db      12 DPFHNEGIPFLV 24
CC
CC      RESULT 24
CC      Q71VB2 PRELIMINARY; PRT; 34 AA.
CC      AC      Q71VB2;
CC      DT      05-JUL-2004 (TrEMBLrel. 27, Created)
CC      DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
CC      DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
CC      DE      Diadenosine triphosphate hydrolase (PHIT protein) (Fragment).
CC      GN      Name=PHIT;
CC      OS      Homo sapiens (Human).
CC      OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC      OX      NCBI_TaxID=9606;
CC      RN      [1]
CC      RP      SEQUENCE FROM N.A.
CC      RC      MEDLINE=98070795; PubMed=9405656;
CC      RA      Inoue H., Ishii H., Alder H., Snyder E., Druck T., Huebner K.,
CC      Croce C.M.;
CC      RT      "Sequence of the FRA3B common fragile region: implications for the
CC      mechanism of PHIT deletion."
CC      RL      Proc. Natl. Acad. Sci. U.S.A. 94:14584-14589(1997).
CC      RN      [2]
CC      RP      SEQUENCE FROM N.A.
CC      RX      MEDLINE=97217778; PubMed=9063739;
CC      RA      Boldog F., Gemmill R.M., West J., Robinson M., Robinson L., Li E.,
CC      Roche J., Todd S., Waggoner B., Lundstrom R., Jacobson J.,
CC      Mulloikandov M.R., Klinger H., Drabkin H.A.;
CC      RT      "Chromosome 3p14 homozygous deletions and sequence analysis of
CC      FRA3B."
CC      RL      Hum. Mol. Genet. 6:193-203(1997).
CC      DR      EMBL; AF020503; AAB97515.1; -.
CC      DR      EMBL; AF023460; AAB84098.1; -.
CC      DR      GO; GO:0016787; P:hydrolase activity; IEA.
CC      KW      Hydrolase.
CC      FT      NON_TER 34
CC      SQ      SEQUENCE 34 AA; 3844 MW; 153C162CDE6C4495 CRC64;

```

DR Pfam;
KW Hydro;

```
Matches 4; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LNSKIAPKIVSQEP 18
    | : : : | : : : |
Db 17 LKTELSPALVNRKP 30

RESULT 29
Q8F047 PRELIMINARY; PRT; 36 AA.
ID Q8F047
AC Q8F047
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LA3650;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu L.-P., Jiang H.-O., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AB011520; AAN50848.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 36 AA; 4421 MW; D43251CF217AB56 CRC64;

Query Match 30.9%; Score 30; DB 2; Length 36;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 NHLNSKIAPKIV 14
    | : | : | : | : |
Db 16 HHRNLKSHFKV 27

RESULT 30
Q7RED1 PRELIMINARY; PRT; 37 AA.
ID Q7RED1
AC Q7RED1
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY05134;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteira M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.B.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.O., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
```

```
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABL01001612; EAA17115.1; -.
DR Hypothetical protein.
SQ SEQUENCE 37 AA; 4243 MW; 1EB3806F42D7C880 CRC64;

Query Match 30.9%; Score 30; DB 2; Length 37;
Best Local Similarity 38.5%; Pred. No. 1.9e+03;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EPHNLNSKIAPKI 13
    | : | : | : | : |
Db 17 DPFHINDGIPFLV 29

RESULT 31
Q817S9 PRELIMINARY; PRT; 37 AA.
ID Q817S9
AC Q817S9
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BC4452;
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapral V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas I.,
RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.C.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis.";
RL Nature 423:87-91(2003).
DR EMBL; AE017012; AAP11365.1; -.
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 4090 MW; A550BD41DE8E9C54 CRC64;

Query Match 30.9%; Score 30; DB 2; Length 37;
Best Local Similarity 40.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 PNLNSKIAPKIVSQ 16
    | : | : | : | : |
Db 5 PSLNLVKFLISVSE 19

RESULT 32
Q8F8V5 PRELIMINARY; PRT; 37 AA.
ID Q8F8V5
AC Q8F8V5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LA0445;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu L.-P., Jiang H.-O., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA
```



```

RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AE011231; AAN47644.1; -.
SQ SEQUENCE 37 AA; 4331 MW; ED56E01F555C5AE5 CRC64;

Query Match      30.9%; Score 30; DB 2; Length 37;
Best Local Similarity 54.5%; Pred. No. 1.9e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      4 HLNKSKIAFKIV 14
DB      11 HKNKKLIFKNV 21

RESULT 33
Q8CR89 PRELIMINARY; PRT; 42 AA.
AC Q8CR89;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein SE11994.
GN OrderedLocusNames=SE11994;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qian Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
DR EMBL; AE016750; AAO05635.1; -.
DR InterPro; IPR008938; ARM.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 42 AA; 5137 MW; DE0BF9C793D77D3 CRC64;

Query Match      30.9%; Score 30; DB 2; Length 42;
Best Local Similarity 46.7%; Pred. No. 2.1e+03;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      2 PNLHNSKIAFKIVSQ 16
DB      23 PNKTRKALKIKQ 37

RESULT 34
Q96BN7 PRELIMINARY; PRT; 43 AA.
AC Q96BN7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; BC015393; AAHL5393.1; -.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 43 AA; 4277 MW; 64A05490EACD7881 CRC64;

Query Match      30.9%; Score 30; DB 2; Length 43;
Best Local Similarity 46.7%; Pred. No. 2.2e+03;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      4 HLNKSKIAFKIVSQEP 18
DB      26 HPNSSIGFQSVPMV 40

RESULT 35
Y576 PYRFU
ID Y576 PYRFU STANDARD; PRT; 46 AA.
AC Q8U394;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical UPF0165 protein PF0576.
GN OrderedLocusNames=PF0576;
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: Belongs to the UPF0165 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE010180; AAL80700.1; -.
DR InterPro; IPR008203; DUF104.
DR Pfam; PF01954; DUF104; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 46 AA; 5442 MW; 517FA67E13A01A6B CRC64;

Query Match      30.9%; Score 30; DB 1; Length 46;
Best Local Similarity 33.3%; Pred. No. 2.3e+03;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      6 NSKIAFKIVSQE 17
DB      24 HSKVIKVIDEE 35

RESULT 36
Q6PKM7 PRELIMINARY; PRT; 46 AA.
AC Q6PKM7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Voltage-dependent L-type Ca2+ channel alpha 1 subunit (Fragment).
GN Name=CACNA1C;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```
RX MEDLINE=21671403; PubMed=11741969;
RA Blumenstein Y., Kanevsky N., Sahar G., Barzilai R., Ivanina T.,
RA Dascal N.;
RT "A novel long N-terminal isoform of human L-type Ca2+ channel is up-
regulated by protein kinase C.";
RL J. Biol. Chem. 277:3419-3423(2002).
[2]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA Blumenstein Y., Kanevsky N., Sahar G., Barzilai R., Ivanina T.,
RA Dascal N.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY604867; AAT02226.1; -.
FT NON TER 46
SQ SEQUENCE 46 AA; 5120 MW; 9AB28AF57F0605CC CRC64;

Query Match 30.9%; Score 30; DB 2; Length 46;
Best Local Similarity 46.2%; Pred. No. 2.3e+03;
Matches 6; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

QY 2 PNHL--NSKIAPK 12
|:|:| |:|:|
Db 17 PSHLSANTEVKFK 29

RESULT 37
AAT02226 PRELIMINARY; PRT; 46 AA.
ID AAT02226
AC AAT02226;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Voltage-dependent L-type Ca2+ channel alpha 1 subunit (Fragment).
GN CACNA1C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21671403; PubMed=11741969;
RA Blumenstein Y., Kanevsky N., Sahar G., Barzilai R., Ivanina T.,
RA Dascal N.;
RT "A novel long N-terminal isoform of human L-type Ca2+ channel is up-
regulated by protein kinase C.";
RL J. Biol. Chem. 277:3419-3423(2002).
[2]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA Blumenstein Y., Kanevsky N., Sahar G., Barzilai R., Ivanina T.,
RA Dascal N.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY604867; AAT02226.1; -.
FT NON TER 46
SQ SEQUENCE 46 AA; 5120 MW; 9AB28AF57F0605CC CRC64;

Query Match 30.9%; Score 30; DB 2; Length 46;
Best Local Similarity 46.2%; Pred. No. 2.3e+03;
Matches 6; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

QY 2 PNHL--NSKIAPK 12
|:|:| |:|:|
Db 17 PSHLSANTEVKFK 29

RESULT 38
Q98156 PRELIMINARY; PRT; 47 AA.
ID Q98156
AC Q98156;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Unidentified ori.
OS Human herpesvirus 8 (HHV-8) (Kaposi's sarcoma-associated herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
```

```
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=37296;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=tumor;
RX MEDLINE=96270595; PubMed=8692871;
RA Zhong W., Wang H., Herndler B., Ganem D.;
RT "Restricted expression of Kaposi sarcoma-associated herpesvirus (human
herpesvirus 8) genes in Kaposi sarcoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:6641-6646(1996).
DR EMBL; U68522; AAC55376.1; -.
DR PIR; JC6158; JC6158.
SQ SEQUENCE 47 AA; 4951 MW; 1EED1CDBBE0455FE CRC64;

Query Match 30.9%; Score 30; DB 2; Length 47;
Best Local Similarity 55.6%; Pred. No. 2.4e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNHLNSKIA 10
|:|:| |:|:|
Db 27 PNHTNGTMA 35

RESULT 39
Q7VKW6 PRELIMINARY; PRT; 30 AA.
ID Q7VKW6
AC Q7VKW6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=HD1746;
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3500CHP / ATCC 700724;
RA Johnson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Munson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017155; AAP96501.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 30 AA; 3581 MW; 125DD8D3598E004 CRC64;

Query Match 30.4%; Score 29.5; DB 2; Length 30;
Best Local Similarity 41.2%; Pred. No. 1.8e+03;
Matches 7; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 1 EPNHLN-SKIAPKIVSQ 16
|:|:| |:|:| |:|:|
Db 5 KPNYLNISTDTFKFLNK 21

RESULT 40
Q8DY47 PRELIMINARY; PRT; 32 AA.
ID Q8DY47
AC Q8DY47;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SAG1646.
GN OrderedLocusNames=SAG1646;
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
```

RA Tettelin H., Massignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Klocir H.M., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V *Streptococcus agalactiae*.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AE014266; AAN00510.1; -;
DR TIGR; SAG1646; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 32 AA; 3831 MW; 2020BF60F88D8AD CRC64;
Query Match 30.4%; Score 29.5; DB 2; Length 32;
Best Local Similarity 46.7%; Pred. No. 2e+03;
Matches 7; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
Qy 6 NSKIAFKIVSQ-EPA 19
Db 16 NNRVKIKIACYEPA 30
::: ||| |||
RESULT 41
Q9KPS8 PRELIMINARY; PRT; 33 AA.
AC Q9KPS8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VC2284.
GN OrderedLocusNames=VC2284;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.B.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004299; AAF95428.1; -;
DR FIR; H82096; H82096.
DR TIGR; VC2284; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 33 AA; 3722 MW; 34A2F0C49EE65E1 CRC64;
Query Match 29.9%; Score 29; DB 2; Length 33;
Best Local Similarity 30.8%; Pred. No. 2.5e+03;
Matches 4; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
Qy 1 EPNHLNSKIAFKI 13
Db 21 DSNQISNDVSFKI 33
: | : : : : :
RESULT 42
Q8F830 PRELIMINARY; PRT; 34 AA.
ID Q8F830;
AC Q8F830;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LA0729;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar Lai;
RX MEDLINE=42598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of *Leptospira*
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AE011261; AAN47928.1; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 34 AA; 3758 MW; 4F1680D378D2F3A CRC64;
Query Match 29.9%; Score 29; DB 2; Length 34;
Best Local Similarity 50.0%; Pred. No. 2.5e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy 6 NSKIAFKIVS 15
Db 12 NVAIAFKVIN 21
||| : : :
RESULT 43
Q9KMU3 PRELIMINARY; PRT; 40 AA.
AC Q9KMU3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VCA0226.
GN OrderedLocusNames=VCA0226;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.B.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004363; AAF96138.1; -;
DR FIR; G82484; G82484.
DR TIGR; VCA0226; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 40 AA; 4543 MW; B7EFC6C4DF62AF28 CRC64;
Query Match 29.9%; Score 29; DB 2; Length 40;
Best Local Similarity 45.5%; Pred. No. 3e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 2 PNHLNSKIAFK 12
||| : : :

Db 30 PNHLELVGGK 40

RESULT 44

Q7QXP2 PRELIMINARY; PRT; 45 AA.
 AC Q7QXP2; 2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE GLP 512 19188 19051.
 OS Giardia lamblia ATCC 50803.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
 OC NCBI_TaxID=184922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D., Olsen G.J., Sogin M.L.;
 RA "Draft sequence of the Giardia lamblia genome.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AACB01000058; EAA39808.1; -;
 DR EMBL; AACB01000058; EAA39808.1; -;
 SQ SEQUENCE 45 AA; 4999 MW; D3BF52CE81B3623D CRC64;

Query Match 29.9%; Score 29; DB 2; Length 45;
 Best Local Similarity 75.0%; Pred. No. 3.4e+03;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 10 AFKIVSQE 17

Db 10 AFKIVSQD 17

RESULT 45

Q8F042 PRELIMINARY; PRT; 46 AA.
 AC Q8F042;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=LA3655;
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OC NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RX MEDLINE=22598143; PubMed=12712204;
 RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H., Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F., Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F., Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W., Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A., Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z., Xu J.-G., Zhao G.-P.;
 RA "Unique physiological and pathogenic features of Leptospira interrogans revealed by whole-genome sequencing.";
 RL Nature 422:888-893(2003).
 DR EMBL; AE011520; AAN50853.1; -;
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 46 AA; 5733 MW; B45CCD35D27AA522 CRC64;

Query Match 29.9%; Score 29; DB 2; Length 46;
 Best Local Similarity 45.5%; Pred. No. 3.5e+03;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 HLNSKIAFKIV 14

Db 22 HQMLSVKFLV 32

RESULT 46

Q72DQ2 PRELIMINARY; PRT; 47 AA.
 AC Q72DQ2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=DVU0877;
 OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
 OC Bacteria; Proteobacteria; Deltaaproteobacteria; Desulfovibrionales; Desulfovibrionaceae; Desulfovibrio.
 OC NCBI_TaxID=882;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=15077118; DOI=10.1038/nb-959;
 RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T., Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J., Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D., Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R., Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RA "The genome sequence of the anaerobic, sulfate-reducing bacterium Desulfovibrio vulgaris Hildenborough.";
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL; AR017312; AAS95357.1; -;
 DR TIGR; DVU0877; -;
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 47 AA; 5214 MW; 49F9FD09C53FE4BB CRC64;

Query Match 29.9%; Score 29; DB 2; Length 47;
 Best Local Similarity 26.7%; Pred. No. 3.6e+03;
 Matches 4; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQEPA 19

Db 1 MSARNTFALTSHEAA 15

RESULT 47

Q7TDN5 PRELIMINARY; PRT; 47 AA.
 AC Q7TDN5;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Halovirus HF1.
 OC Viruses; unclassified viruses; Haloviruses.
 OC NCBI_TaxID=222645;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=15090523;
 RA Tang S.-L., Nurtall S., Dyal-Smith M.;
 RA "Haloviruses HF1 and HF2: Evidence for a Recent and Large
 RT Recombination Event.";
 RL J. Bacteriol. 186:2810-2817(2004).
 DR EMBL; AY190604; AAO61306.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 47 AA; 5279 MW; FDDF8C11C54306BE CRC64;

Query Match 29.9%; Score 29; DB 2; Length 47;
 Best Local Similarity 31.6%; Pred. No. 3.6e+03;
 Matches 6; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSQEPA 19

Db 10 EPVKNKEATFIIIESFPS 28

```

RESULT 48
Q8V6V0
ID Q8V6V0 PRELIMINARY; PRT; 47 AA.
AC Q8V6V0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Halovirus HF2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=33771;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21964609; PubMed=11967086;
RA Tang S.-L., Nuttall S., Ngui K., Fisher C., Lopez P., Dyll-Smith M.;
RT "HP2: a double-stranded DNA tailed haloarchaeal virus with a mosaic
genome.";
RL Mol. Microbiol. 44:283-296(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX Tang S.-L., Fisher C., Ngui K., Nuttall S.D., Dyll-Smith M.L.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF222060; AAL54929.1; -.
DR EMBL; AF222060; AAL54929.1; -.
KW Hypothetical protein.
SQ SEQUENCE 47 AA; 5279 MW; FPDF8C11C54306BE CRC64;

Query Match 29.9%; Score 29; DB 2; Length 47;
Best Local Similarity 31.6%; Pred. No. 3.6e+03;
Matches 6; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSQEPA 19
||| : : : : :
Db 10 EPVVKGCEATFEIESFES 28

RESULT 49
AAS95357
ID AAS95357 PRELIMINARY; PRT; 47 AA.
AC AAS95357;
DT 27-APR-2004 (TrEMBLrel. 27, Created)
DT 27-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN DVU0877.
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118;
RA Heidelberg J.F., Seshadri R., Havenan S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
RA Doolittle R.F., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Daviden T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AB017312; AAS95357.1; -.
DR TIGR; DVU0877; -.
KW Hypothetical protein.
SQ SEQUENCE 47 AA; 5214 MW; 49F9FD0C53FE4BB CRC64;

Query Match 29.9%; Score 29; DB 2; Length 47;
Best Local Similarity 26.7%; Pred. No. 3.6e+03;
Matches 4; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQEPA 19

```

```

Db 1 MSARVTFELTSHEAA 15
::: : : : :

RESULT 50
Q95N96
ID Q95N96 PRELIMINARY; PRT; 48 AA.
AC Q95N96;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 2,4-dienoyl-CoA reductase (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22042656; PubMed=12047237;
RA Clop A., Cercos A., Tomas A., Perez-Enciso M., Varona L.,
RA Noguera J.L., Sanchez A., Amills M.;
RT "Assignment of the 2,4-dienoyl-CoA reductase (DECR) gene to porcine
chromosome 4.";
RL Anim. Genet. 33:164-165(2002).
DR EMBL; AF335499; AAK52724.1; -.
FT NON_TER 1
FT NON_TER 48
SQ SEQUENCE 48 AA; 5215 MW; 2F6E7D85C40EA56C CRC64;

Query Match 29.9%; Score 29; DB 2; Length 48;
Best Local Similarity 50.0%; Pred. No. 3.6e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PNHLSKIAF 11
||| : : : :
Db 23 PNTFQGVAF 32

RESULT 51
Q95418
ID Q95418 PRELIMINARY; PRT; 49 AA.
AC Q95418;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Madan A., Rowen L., Qin S., Dickhoff R., Shaffer T., James R.,
RA Abbasi N., Loretz C., Madan A., Dors M., Dahl T., Hall J., Lasky S.,
RA Hood L.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF107885; AAC79728.1; -.
KW Hypothetical protein.
FT NON_TER 49
SQ SEQUENCE 49 AA; 5479 MW; C5D1658A74CB294F CRC64;

Query Match 29.9%; Score 29; DB 2; Length 49;
Best Local Similarity 83.3%; Pred. No. 3.7e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NHLNGK 8
||| : : :
Db 35 NHLNGK 40

RESULT 52
Q7YS12
ID Q7YS12 PRELIMINARY; PRT; 50 AA.

```

AC	Q7YS12;	OS	Caulobacter crescentus.	OS	Caulobacter crescentus.
DT	01-OCT-2003 (TrEMBLrel. 25, Created)	OC	Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;	OC	Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)	OC	Caulobacteriaceae; Caulobacter.	OC	Caulobacteriaceae; Caulobacter.
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	OX	NCBI_TaxID=155892;	OX	NCBI_TaxID=155892;
DE	RYR1 protein (Fragment).	[1]		[1]	
GN	Name=RYR1;	RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
OS	Ovis aries (Sheep).	RC	STRAIN=ATCC 19089 / CB15;	RC	STRAIN=ATCC 19089 / CB15;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	RX	MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;	RX	MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	RA	Nierman W.C., Reibdyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,	RA	Nierman W.C., Reibdyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
OC	Caprinae; Ovis.	RA	Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,	RA	Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
OC	NCBI_TaxID=9940;	RA	Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,	RA	Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
OX		RA	DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hart D.H.,	RA	DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hart D.H.,
RP	SEQUENCE FROM N.A.	RA	Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,	RA	Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA	Li X.L.;	RA	Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,	RA	Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.	RA	Ermoiaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,	RA	Ermoiaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
DR	EMBL; AY291439; AAP72294.1; -.	RA	Fraser C.M.;	RA	Fraser C.M.;
FT	NON_TER 1 1	RL	"Complete genome sequence of Caulobacter crescentus.";	RL	"Complete genome sequence of Caulobacter crescentus.";
FT	NON_TER 50 50	RL	Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).	RL	Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
SQ	SEQUENCE 50 AA; 5534 MW; 3423DB673886C9DC CRC64;	DR	EMBL; AE005797; AAK23180.1; -.	DR	EMBL; AE005797; AAK23180.1; -.
		DR	PIR; H87397; H87397.	DR	PIR; H87397; H87397.
		DR	TIGR; CC1197; -.	DR	TIGR; CC1197; -.
		KW	Complete proteome; Hypothetical protein.	KW	Complete proteome; Hypothetical protein.
		SQ	SEQUENCE 49 AA; 5145 MW; 54B3C821D21A6AC0 CRC64;	SQ	SEQUENCE 49 AA; 5145 MW; 54B3C821D21A6AC0 CRC64;
Qy	5 LNSKIAFKIVSOEA 19	Query Match	29.9%; Score 29; DB 2; Length 50;	Query Match	29.4%; Score 28.5; DB 2; Length 49;
		Best Local Similarity	40.0%; Pred. No. 3.8e+03;	Best Local Similarity	50.0%; Pred. No. 4.5e+03;
		Matches	6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;	Matches	7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;
Db	3 MTDXLAFDVGLQEDA 17				
		Qy	3 NLSKIAFKIVSOEA 15	Qy	3 NLSKIAFKIVSOEA 15
			: :		: :
		Db	18 NLSKIAFKIVSOEA 31	Db	18 NLSKIAFKIVSOEA 31
RESULT 53				RESULT 55	
Q8XMB6		Q9AJQ7	PRELIMINARY; PRT; 50 AA.	Q9AJQ7	PRELIMINARY; PRT; 23 AA.
ID	Q8XMB6	ID	Q9AJQ7	ID	Q9AJQ7
AC	Q8XMB6;	AC	Q9AJQ7;	AC	Q9AJQ7;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)	DT	01-JUN-2001 (TrEMBLrel. 17, Created)	DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)	DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Hypothetical protein CPE0773.	DE	Putative aspartokinase III (Fragment).	DE	Putative aspartokinase III (Fragment).
GN	OrderedLocusNames=CPE0773;	GN	Name=lysc;	GN	Name=lysc;
OS	Clostridium perfringens.	OS	Vibrio fischeri.	OS	Vibrio fischeri.
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;	OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;	OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Clostridium.	OC	Vibrionaceae; Vibrio.	OC	Vibrionaceae; Vibrio.
OX	NCBI_TaxID=1502;	OX	NCBI_TaxID=668;	OX	NCBI_TaxID=668;
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=13;	RC	STRAIN=ATCC 7744;	RC	STRAIN=ATCC 7744;
RX	MEDLINE=21664373; PubMed=11792842;	RX	MEDLINE=21147939; PubMed=11250084;	RX	MEDLINE=21147939; PubMed=11250084;
RA	Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,	RA	Kasai S., Yamazaki T.;	RA	Kasai S., Yamazaki T.;
RA	Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;	RA	"Identification of the cobalamin-dependent methionine synthase gene,	RA	"Identification of the cobalamin-dependent methionine synthase gene,
RT	"Complete genome sequence of Clostridium perfringens, an anaerobic	RT	meth, in Vibrio fischeri ATCC 7744 by sequencing using genomic DNA as	RT	meth, in Vibrio fischeri ATCC 7744 by sequencing using genomic DNA as
RT	flesh-eater.";	RT	a template.";	RT	a template.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).	RL	Gene 264:281-288(2001).	RL	Gene 264:281-288(2001).
DR	EMBL; AP003188; BAB80479.1; -.	DR	EMBL; AB039955; BAB39356.1; -.	DR	EMBL; AB039955; BAB39356.1; -.
KW	Complete proteome; Hypothetical protein.	DR	GO: GO:0016301; P:kinase activity; IEA.	DR	GO: GO:0016301; P:kinase activity; IEA.
SQ	SEQUENCE 50 AA; 5894 MW; EA6C276FEC59D2F5 CRC64;	KW	Kinase.	KW	Kinase.
		FT	NON_TER 1 1	FT	NON_TER 1 1
		SQ	SEQUENCE 23 AA; 2766 MW; 950DC44A393C51E7 CRC64;	SQ	SEQUENCE 23 AA; 2766 MW; 950DC44A393C51E7 CRC64;
		Query Match	29.9%; Score 29; DB 2; Length 50;	Query Match	28.9%; Score 28; DB 2; Length 23;
		Best Local Similarity	50.0%; Pred. No. 3.8e+03;	Best Local Similarity	53.8%; Pred. No. 2.5e+03;
		Matches	6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	Matches	7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy	3 NLSKIAFKIV 14				
	: :				
Db	32 NLSKIAFKIV 43				
RESULT 54				RESULT 56	
Q9A8Z8		Q01228	PRELIMINARY; PRT; 49 AA.	Q01228	PRELIMINARY; PRT; 24 AA.
ID	Q9A8Z8	ID	Q01228	ID	Q01228
AC	Q9A8Z8;	AC	Q01228;	AC	Q01228;
DT	01-JUN-2001 (TrEMBLrel. 17, Created)	DT	01-JUN-2001 (TrEMBLrel. 17, Created)	DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Hypothetical protein CC1197.	DE	Hypothetical protein CC1197.	DE	Hypothetical protein CC1197.
GN	OrderedLocusNames=CC1197;	GN	OrderedLocusNames=CC1197;	GN	OrderedLocusNames=CC1197;

AC	Q01228;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DI	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	ORF1 protein (Fragment).
GN	Name=ORF1;
OS	Williopsis saturnus var. suaveolens.
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Williopsis.
OX	NCBI_TaxID=58637;
[1]	
RN	SEQUENCE 24 AA; 2877 MW; 5236F1C0BC46585F CRC64;
RP	SEQUENCE FROM N.A.
RC	STRAIN=CBS 255;
RA	Nosek J.;
RL	Submitted (FEB-1994) to the EMBL/GenBank/DBDJ databases.
DR	EMBL; X77691; CAA54773.1; -.
DR	PIR; S47281; S47281.
FT	NON TER 24
SQ	SEQUENCE 24 AA; 2877 MW; 5236F1C0BC46585F CRC64;
Query Match	Score 28; DB 2; Length 24;
Best Local Similarity	36.4%; Pred. No. 2.6e+03;
Matches	5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY	3 HLNSKIAFKIVS 15
Db	::: :::
	12 SNTNNKILFRLES 24
RESULT 57	
SPIG PSEUS	
ID _SPIG PSEUS	STANDARD; PRT; 25 AA.
AC	P82357;
DT	16-OCT-2001 (Rel. 40, Created)
DI	16-OCT-2001 (Rel. 40, Last sequence update)
DE	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Spingerin [Contains: Spingerin N-3; Spingerin C-4].
OS	Pseudocanthotermes spiniger.
OC	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC	Neoptera; Orthopteroidea; Dictyoptera; Isoptera; Termitidae;
OC	Macrotermitinae; Pseudocanthotermes.
OX	NCBI_TaxID=115113;
[1]	
RP	SEQUENCE, MASS SPECTROMETRY, AND FUNCTION.
RC	TISSUE=Blood, and Salivary gland;
RX	PubMed=11053427; DOI=10.1074/jbc.M002998200;
RA	Lamberty M., Zachary D., Lanot R., Bordereau C., Robert A.,
RA	Hoffmann J.A., Bulet P.;
RT	"Insect immunity. Constitutive expression of a cysteine-rich
RT	antifungal and a linear antibacterial peptide in a termite insect.";
RL	J. Biol Chem. 276:4085-4092(2001).
-1-	FUNCTION: Active against Gram-positive bacteria B.megaterium and
CC	M.luteus, Gram-negative bacteria E.coli SBS363 and D22,
CC	K.pneumoniae, S.typhimurium and P.aeruginosa, Yeast C.albicans and
CC	filamentous fungi F.culmorum, N.crassa, N.hematococca and
CC	T.pyridae. Inactive against Gram-positive bacteria B.subtilis,
CC	S.pyogenes, B.thuringiensis and S.aureus, Gram-negative bacteria
CC	E.coliaceae and E.carotovora and filamentous fungus B.bassiana.
-1-	SUBCELLULAR LOCATION: Secreted.
CC	-1- INDUCTION: By bacterial infection.
CC	-1- MASS SPECTROMETRY: MW=3001.8; METHOD=MALDI; RANGE=1-25;
CC	NOTE=Ref.1.
KW	Antibiotic; Direct protein sequencing; Fungicide.
FT	CHAIN 1 25 Spingerin.
FT	CHAIN 4 25 Spingerin N-3.
FT	CHAIN 1 21 Spingerin C-4.
SQ	SEQUENCE 25 AA; 3001 MW; AA79370264262F60 CRC64;
Query Match	Score 28; DB 1; Length 25;
Best Local Similarity	36.4%; Pred. No. 2.7e+03;
Matches	4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY	4 HLNSKIAFKIV 14

Q9TTH7
ID Q9TTH7 PRELIMINARY; PRT; 37 AA.
AC Q9TTH7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Skeletal muscle-specific calpain (fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Nonnenan D., Kochmarai M.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087571; AAF23263.1; -.
DR HSP; Q07009; IDFO.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004198; F:calpain activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001300; Peptidase_C2.
DR Pfam; PF00648; Peptidase_C2; 1.
DR PROSITE; PS50203; CALPAIN_CAT; 1.
FT NON TER 1
FT NON TER 37
SQ SEQUENCE 37 AA; 4325 MW; CEF3BD31015508BF CRC64;
Query Match 28.9%; Score 28; DB 2; Length 37;
Best Local Similarity 23.1%; Pred. No. 4.1e+03;
Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 5 LNSKIAFKIVSQE 17
Db 13 LNKRLFRVIPHD 25
RESULT 63
Q7PH38
ID Q7PH38 PRELIMINARY; PRT; 38 AA.
AC Q7PH38;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000023653.
GN Name=ENSANG00000021015;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA501008880; EAA44699.1; -.
SQ SEQUENCE 38 AA; 4528 MW; 845520B4B0A797CB CRC64;
Query Match 28.9%; Score 28; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 4.2e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 EPNHLSKIAPK 12
Db 14 QPNFRFSEAFK 25
RESULT 64
Q8CRL7
ID Q8CRL7 PRELIMINARY; PRT; 41 AA.

QY 4 HLNSKIAFKIVSQEP 18
Db 20 HILQMKNKLSQNP 34
RESULT 60
Q71MZ3
ID Q71MZ3 PRELIMINARY; PRT; 35 AA.
AC Q71MZ3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein su2 (fragment).
GN Name=su2;
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RA Su L., Wu B., Li F., Li W., Sun H.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF434671; AAQ04270.1; -.
KW Hypothetical protein.
FT NON TER 35
FT NON TER 35
SQ SEQUENCE 35 AA; 4282 MW; 02D24ECA5193209F CRC64;
Query Match 28.9%; Score 28; DB 2; Length 35;
Best Local Similarity 71.4%; Pred. No. 3.9e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 3 NHLNSKI 9
Db 16 NHLDSSEI 22
RESULT 61
AAQ04270
ID AAQ04270 PRELIMINARY; PRT; 35 AA.
AC AAQ04270;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein su2 (fragment).
GN SU2.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715;
RA Su L., Wu B., Li F., Li W., Sun H.;
RL "Cloning and sequence analysis of cell division protein ftsK of
RL Streptococcus sobrinus 6715."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF434671; AAQ04270.1; -.
KW Hypothetical protein.
FT NON TER 35
FT NON TER 35
SQ SEQUENCE 35 AA; 4282 MW; 02D24ECA5193209F CRC64;
Query Match 28.9%; Score 28; DB 2; Length 35;
Best Local Similarity 71.4%; Pred. No. 3.9e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 3 NHLNSKI 9
Db 16 NHLDSSEI 22
RESULT 62


```

AC Q8CRL7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SE1741.
GN OrderedLocusNames=SE1741;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qiu Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
DR EMBL; AB016749; AA05340.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 41 AA; 5034 MW; E7C4BBA0059071CD CRC64;

Query Match 28.9%; Score 28; DB 2; Length 41;
Best Local Similarity 55.6%; Pred. No. 4.6e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 HLNSKIAPK 12
Db 32 HSLKVDK 40

RESULT 65
Q8EZK3
ID Q8EZK3 PRELIMINARY; PRT; 41 AA.
AC Q8EZK3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LA3850;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AB011540; AA51048.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 41 AA; 4780 MW; 308AC1BESAA3F597 CRC64;

Query Match 28.9%; Score 28; DB 2; Length 41;
Best Local Similarity 55.6%; Pred. No. 4.6e+03;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 KIAFKIVSQ 16
Db 17 KVAFTVSK 25

RESULT 66
Q7TDL0

```

```

ID Q7TDL0 PRELIMINARY; PRT; 43 AA.
AC Q7TDL0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Halovirus HF1.
OC Viruses; unclassified viruses; Haloviruses.
OX NCBI_TaxID=222645;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15090523;
RA Tang S.-L., Nuttall S., Dyll-Smith M.;
RT "Haloviruses HF1 and HF2: Evidence for a Recent and Large
RT Recombination Event.";
RL J. Bacteriol. 186:2810-2817(2004).
DR EMBL; AY190604; AA061331.1; -.
KW Hypothetical protein.
SQ SEQUENCE 43 AA; 4285 MW; CB3F8476D8EALC74 CRC64;

Query Match 28.9%; Score 28; DB 2; Length 43;
Best Local Similarity 55.6%; Pred. No. 4.8e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKI 9
Db 31 EPNHVVNPI 39

RESULT 67
Q8V6S1
ID Q8V6S1 PRELIMINARY; PRT; 43 AA.
AC Q8V6S1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Halovirus HF2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=33771;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21964609; PubMed=11967086;
RA Tang S.-L., Nuttall S., Ngui K., Fisher C., Lopez P., Dyll-Smith M.;
RT "HF2: a double-stranded DNA tailed haloarchaeal virus with a mosaic
RT genome.";
RL Mol. Microbiol. 44:283-296(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Tang S.-L., Fisher C., Ngui K., Nuttall S.D., Dyll-Smith M.L.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF222060; AAL54954.1; -.
KW Hypothetical protein.
SQ SEQUENCE 43 AA; 4285 MW; CB3F8476D8EALC74 CRC64;

Query Match 28.9%; Score 28; DB 2; Length 43;
Best Local Similarity 55.6%; Pred. No. 4.8e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKI 9
Db 31 EPNHVVNPI 39

RESULT 68
Q8VE64
ID Q8VE64 PRELIMINARY; PRT; 46 AA.
AC Q8VE64;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).

```

RC	STRAIN=CAVRS;
RX	MEDLINE=20416036; PubMed=10961854;
RA	Le Jambre L.F., Lenane I.J., Wardrop A.J.;
RT	"A hybridization technique to identify anthelmintic resistance genes
RR	in Haemonchus.";
RL	Int. J. Parasitol. 29:1979-1985 (1999).
DR	EMBL; AF182012; AAF03708.1; -.
FT	NON_TER 1 1
FT	NON_TER 49 49
SQ	SEQUENCE 49 AA; 5671 MW; 7C91E57B80C529B8 CRC64;
Query Match 28.9%; Score 28; DB 2; Length 49;	
Best Local Similarity 83.3%; Pred. No. 5.5e+03;	
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	13 IVSQEP 18
DB	:
	26 VVSQEP 31
 RESULT 71	
Q9U6N9	PRELIMINARY; PRT; 49 AA.
ID Q9U6N9	PRELIMINARY; PRT; 49 AA.
AC Q9U6N9	01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000	(TReMBLrel. 13, Last sequence update)
DT 01-MAR-2002	(TReMBLrel. 20, Last annotation update)
DE	P-glycoprotein 1 (Fragment).
GN	Name=pgp-1;
OS	Haemonchus contortus (Barber pole worm).
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC	Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
NBI_TaxID=6289;	[1]
RN	SEQUENCE FROM N.A.
PC	STRAIN=CAVRS;
RX	MEDLINE=20416036; PubMed=10961854;
RA	Le Jambre L.F., Lenane I.J., Wardrop A.J.;
RT	"A hybridization technique to identify anthelmintic resistance genes
RR	in Haemonchus.";
RL	Int. J. Parasitol. 29:1979-1985 (1999).
DR	EMBL; AF182011; AAF03707.1; -.
FT	NON_TER 1 1
FT	NON_TER 49 49
SQ	SEQUENCE 49 AA; 5687 MW; D80F8A7E80C529B2 CRC64;
Query Match 28.9%; Score 28; DB 2; Length 49;	
Best Local Similarity 83.3%; Pred. No. 5.5e+03;	
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	13 IVSQEP 18
DB	:
	26 VVSQEP 31
 RESULT 72	
Q7MBK7	PRELIMINARY; PRT; 49 AA.
ID Q7MBK7	PRELIMINARY; PRT; 49 AA.
AC Q7MBK7	01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004	(TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004	(TReMBLrel. 26, Last annotation update)
DE	Hypothetical protein vvp36.
GN	Name=VVP36;
OS	Vibrio vulnificus (strain YJ016).
OG	Plasmid pYJ016.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Vibrionaceae; Vibrio.
NCBI_TaxID=196600;	[1]
RN	SEQUENCE FROM N.A.
RP	PubMed=14565965;
RX	Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,

```

RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; AP005352; BAC97759.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 49 AA; 5690 MW; 9AA8C036F84PADB4 CRC64;

Query Match      28.9%; Score 28; DB 2; Length 49;
Best Local Similarity 36.4%; Pred. No. 5.5e+03;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 PNHLNKTAFK 12
Db 36 PRHKNTLVLR 46

RESULT 73
Q8CLV6      28 AA; 3166 MW; B98B099C1B1B65C6 CRC64;
ID Q8CLV6      PRELIMINARY; PRT; 28 AA.
AC Q8CLV6
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical.
GN OrderedLocusNames=y0121;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersiniaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AE013612; AAM83715.1; -.
KW Hypothetical protein.
SQ SEQUENCE 28 AA; 3166 MW; B98B099C1B1B65C6 CRC64;

Query Match      28.4%; Score 27.5; DB 2; Length 28;
Best Local Similarity 40.0%; Pred. No. 3.7e+03;
Matches 8; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 2 PNHLNKTAFKIVSQEP 18
Db 7 PNHLNKTAFKIVSQEP 26

RESULT 74
Q8FZX9      30 AA; 30 AA.
ID Q8FZX9      PRELIMINARY; PRT; 30 AA.
AC Q8FZX9
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BR1341;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,

```

```

RA Read T.D., Dodson R.J., Unyam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014431; AAN30255.1; -.
DR TIGR; BR1341; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 30 AA; 3390 MW; D81F2937D888009B CRC64;

Query Match      28.4%; Score 27.5; DB 2; Length 30;
Best Local Similarity 38.1%; Pred. No. 4e+03;
Matches 8; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 1 EPNHLNSK--IAFKIVSQEP 18
Db 10 EPAHANGRERMIAPKLOQLRP 30

RESULT 75
Q7RMJ0      36 AA.
ID Q7RMJ0      PRELIMINARY; PRT; 36 AA.
AC Q7RMJ0
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
GN Name=PY02190;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XN1;
RX PubMed=12368865;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perlea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA Van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL0100600; EAA21623.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 36 AA; 4364 MW; 7DC69EAAAC714A14 CRC64;

Query Match      28.4%; Score 27.5; DB 2; Length 36;
Best Local Similarity 36.8%; Pred. No. 4.9e+03;
Matches 7; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 1 EPNHLNSKIAFKIVSQEPA 19
Db 3 EPTHL-----FSMYDEETA 16

RESULT 76
Q9Y565      PRELIMINARY; PRT; 25 AA.
ID Q9Y565
AC Q9Y565;

```

```

RC STRAIN=NRRL# Y-12651;
RA Green-Willms N.S., Costanzo M.C., Fox T.D.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U83662; AAB41431.1; -.
FT NON TER 29
SQ SEQUENCE 29 AA; 3171 MW; 993D9BCEFFD0C012 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 29;
Best Local Similarity 44.4%; Pred. No. 4.7e+03;
Matches 8; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 3 NHL--NSKIAFKIVSQEP 18
   :|||:|||:|:|
Db 6 SHLLNSRIAHVPSKKP 23

RESULT 79
Q8WD36 PRELIMINARY; PRT; 31 AA.
ID Q8WD36;
AC Q8WD36;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NADH dehydrogenase subunit 5 (Fragment).
GN Name=ND5;
OS Melipona bicolor.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Meliponina.
OX NCBI_TaxID=60889;
RN [1]
RP SEQUENCE FROM N.A.
RA Silvestre D., Francisco F.O., Weinlich R., Arias M.C.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF370441; AAJ57238.1; -.
GO GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 1
SQ SEQUENCE 31 AA; 3856 MW; 1FC18081D4D0F53F CRC64;

Query Match 27.8%; Score 27; DB 2; Length 31;
Best Local Similarity 45.5%; Pred. No. 5.1e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVS 15
   :|||:|||:|
Db 3 MNXKIKFNSIS 13

RESULT 80
RL26_XENLA STANDARD; PRT; 33 AA.
ID RL26_XENLA;
AC P49629;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 60S ribosomal protein L26 (Fragment).
GN Name=RPL26;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92253404; PubMed=1579486;
RX Loreni F., Francesconi A., Jappelli R., Amaldi F.;
RL "Analysis of mRNAs under translational control during Xenopus
RT embryogenesis: isolation of new ribosomal protein clones."
RL Nucleic Acids Res. 20:1859-1863(1992).
CC -!- SIMILARITY: Belongs to the L24p family of ribosomal proteins.
-----

```

```

DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Canalicular multispecific organic anion-transporter (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hu L.J., Simon F.R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF163577; AAD43590.1; -.
FT NON TER 25
SQ SEQUENCE 25 AA; 3018 MW; 75C5187F0B630BAC CRC64;

Query Match 27.8%; Score 27; DB 2; Length 25;
Best Local Similarity 33.3%; Pred. No. 4e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 7; Gaps 0;

QY 3 NHLNSKIAFKIVSQE 17
   |||:|:|
Db 10 NHINERFVDPVQGE 24

RESULT 77
Q9UWG8 PRELIMINARY; PRT; 27 AA.
ID Q9UWG8;
AC Q9UWG8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE keto-valine-ferredoxin oxidoreductase delta-subunit, VOR delta
DE (Fragment).
OS Thermococcus sp.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus
OX NCBI_TaxID=35749;
RN [1]
RP SEQUENCE.
RC STRAIN=ES-1;
RX MEDLINE=96146528; PubMed=8550513;
RA Heider J., Mai X., Adams M.W.;
RT "Characterization of 2-ketoisovalerate ferredoxin oxidoreductase, a
RT new and reversible coenzyme A-dependent enzyme involved in peptide
RT fermentation by hyperthermophilic archaea."
RL J. Bacteriol. 178:780-787(1996).
SQ SEQUENCE 27 AA; 3046 MW; 60C56F313C76E16C CRC64;

Query Match 27.8%; Score 27; DB 2; Length 27;
Best Local Similarity 45.5%; Pred. No. 4.4e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 KIAFKIVSQEP 18
   :|||:|||:|
Db 14 KLIVFASVDEYP 24

RESULT 78
P78747 PRELIMINARY; PRT; 29 AA.
ID P78747;
AC P78747;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE K-MRP51 (Fragment).
GN Name=k-MRP51;
OS Saccharomyces kluyveri (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4934;
RN [1]
RP SEQUENCE FROM N.A.
-----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X64211; CAB56811.1; -;
 DR PIR; S22603; S22603.
 DR InterPro; IPR005825; Ribosomal_L24_26.
 DR InterPro; IPR008991; Transl_SH3-like.
 DR PROSITE; PS01108; RIBOSOMAL_L24; PARTIAL.
 KW Ribosomal protein.
 FT NON_TER 1
 FT NON_TER 33

SQ SEQUENCE 33 AA; 4055 MW; DE5DB3E255B0BA92 CRC64;
 Query Match 27.8%; Score 27; DB 1; Length 33;
 Best Local Similarity 37.5%; Pred. No. 5.4e+03;
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 PNHLNSKIAPKIVSQE 17
 Db 1 PSHVRRKIMGWPLSKE 16

RESULT 81

Q71NE3
 ID Q71NE3 PRELIMINARY; PRT; 33 AA.
 AC Q71NE3
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Photosystem II T protein (Fragment).
 GN Namespsb;
 GN Oxyrrhynchium savatieri.
 OS Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 OC Bryopsida; Bryidae; Hypnales; Brachytheciaceae; Oxyrrhynchium.
 OX NCBI_TaxID=184658;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Huttunen S., Ignatov M.S.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF417453; AAQ03937.1; -;
 DR GO; GO:0009507; Chloroplast; IEA.
 DR InterPro; IPR001743; PSII_PsbT.
 DR Pfam; PF01405; PsbT; 1.
 DR ProDom; PD004453; PSII_PsbT; 1.
 KW Chloroplast.
 FT NON_TER 1
 FT NON_TER 33

SQ SEQUENCE 33 AA; 3697 MW; 96A76F1A1B53D0FF CRC64;
 Query Match 27.8%; Score 27; DB 2; Length 33;
 Best Local Similarity 60.0%; Pred. No. 5.4e+03;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
 Db 15 IFFAIFSRP 24

RESULT 82

AAQ03937
 ID AAQ03937 PRELIMINARY; PRT; 33 AA.
 AC AAQ03937;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Photosystem II T protein (Fragment).
 GN PSBT.
 OS Oxyrrhynchium savatieri.

OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 OC Bryopsida; Bryidae; Hypnales; Brachytheciaceae; Oxyrrhynchium.
 OX NCBI_TaxID=184658;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Huttunen S., Ignatov M.S.;
 RL "Phylogeny of moss family Brachytheciaceae based on morphological and
 RL molecular trnL-F, ITS2 and psbT-H data."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF417453; AAQ03937.1; -;
 KW Chloroplast; Photosystem II.
 FT NON_TER 1
 FT NON_TER 33

SQ SEQUENCE 33 AA; 3697 MW; 96A76F1A1B53D0FF CRC64;
 Query Match 27.8%; Score 27; DB 2; Length 33;
 Best Local Similarity 60.0%; Pred. No. 5.4e+03;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
 Db 15 IFFAIFSRP 24

RESULT 83

Q8WY57
 ID Q8WY57 PRELIMINARY; PRT; 34 AA.
 AC Q8WY57;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Epithelial sodium channel beta-2 subunit (Fragment).
 GN Name-SCNN1B;
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas C.P., Loftus R.W., Liu K.Z., Itani O.A.;
 RL "Genomic organization of the 5' end of human beta-ENaC and preliminary
 RL characterization of its promoter."
 RL Am. J. Physiol. Renal Physiol. 282:F898-F909 (2002).
 DR EMBL; AF260228; AAL48196.1; -;
 DR GO; GO:0005216; Filon channel activity; IEA.
 KW Ionic channel.
 FT NON_TER 34
 FT NON_TER 34

SQ SEQUENCE 34 AA; 3940 MW; 5A703247B3A73FE7 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 34;
 Best Local Similarity 33.3%; Pred. No. 5.6e+03;
 Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 HLNSKIAFKIVSQEP 18
 Db 4 HINPAYFLKLHGFP 18

RESULT 84

Q7M4R4
 ID Q7M4R4 PRELIMINARY; PRT; 36 AA.
 AC Q7M4R4;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Fibroblast-activating factor 32K precursor (fragments).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.

```

RX MEDLINE=91224664; PubMed=2026444;
RA Demeter J., Medzhiradzky D., Kna H., Goetzl E.J., Turk C.W.;
RT "Isolation and partial characterization of the structures of
RT fibroblast activating factor-related proteins from U937 cells.";
RL Immunology 72:350-354 (1991).
DR PIR; A61235; A61235.
FT NON_TER 1
FT NON_TER 36
FT NON_TER 36
SQ SEQUENCE 36 AA; 4246 MW; EE50DBCA4BE0274 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 36;
Best Local Similarity 33.3%; Pred. No. 5.9e+03;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 6 NSKIAFKIV 14
| : : : :
Db 24 NDRVAYKVL 32

RESULT 85
Q9RR54 PRELIMINARY; PRT; 36 AA.
AC Q9RR54; 07CR59;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Invasol SirA.
GN Name=sirA; OrderedLocusNames=STM0274.1; ORFNames=STM0274A;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=82/6915;
RX MEDLINE=97366785; PubMed=9223615;
RA Park J.U.;
RT "Molecular analysis of the genes mediating Salmonella invasion.";
RL FEMS Immunol. Med. Microbiol. 18:113-117 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
DR EMBL; AF026035; AAF08680.1; -.
DR EMBL; AE008707; AAL19231.1; -.
KW Complete proteome.
SQ SEQUENCE 36 AA; 4143 MW; 2PC5E996AD27EBDCD CRC64;

Query Match 27.8%; Score 27; DB 2; Length 36;
Best Local Similarity 46.2%; Pred. No. 5.9e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 PNHLNSKIAFKIV 14
| : : : :
Db 23 PNASNNAEFYII 35

RESULT 86
Q87854 PRELIMINARY; PRT; 37 AA.
AC Q87854;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Sps0976.

```

```

GN OrderedLocusNames=SPS0976;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1;
RX MEDLINE=22683278; PubMed=12799345;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055 (2003).
DR EMBL; AP005144; BAC64071.1; -.
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 4210 MW; 917044DB38A5C045 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 37;
Best Local Similarity 66.7%; Pred. No. 6.1e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 KIAFKIVSQ 16
| : : : :
Db 12 KIPLKIVAQ 20

RESULT 87
Q8P140 PRELIMINARY; PRT; 37 AA.
AC Q8P140;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein spym18_1077.
GN OrderedLocusNames=spym18_1077;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gorpel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith I.W., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AE010034; AAL97700.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 37 AA; 4209 MW; 917044DB38AB20AB CRC64;

Query Match 27.8%; Score 27; DB 2; Length 37;
Best Local Similarity 66.7%; Pred. No. 6.1e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 KIAFKIVSQ 16
| : : : :
Db 12 KIPLKIVAQ 20

RESULT 88
Y520 BORBU
ID Y520 BORBU STANDARD; PRT; 38 AA.
AC O51470;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

```

01-OCT-2004 (Rel. 45, Last annotation update)
DT Hypothetical protein BB0520.
GN OrderedLocusNames=BB0520;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Castlens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Winn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Uterback T.R., Matthey L., McDonald L.A.,
RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.";
RL Nature 390:580-586(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001154; AAC66896.1; -
DR PIR; G70164; G70164.
DR TIGR; BB0520; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 38 AA; 4574 MW; E7EC32520237F622 CRC64;

Query Match 27.8%; Score 27; DB 1; Length 38;
Best Local Similarity 38.5%; Pred. No. 6.3e+03;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSOE 17
Db :||| |:::
20 INKKIIFLTKK 32

RESULT 89
Q8TGN0 PRELIMINARY; PRT; 38 AA.
ID Q8TGN0
AC Q8TGN0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein YKL156C-A.
GN Name=YKL156C-A;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21624570; PubMed=11753363;
RA Kumar A., Harrison P.M., Cheung K.H., Lan N., Echols N., Bertone P.,
RA Miller P., Gerstein M.B., Snyder M.;
RT "An integrated approach for finding overlooked genes in yeast.";
RL Nat. Biotechnol. 20:58-63(2002).
DR EMBL; AF479960; AL479273.1; -
KW Hypothetical protein.
SQ SEQUENCE 38 AA; 4232 MW; A26E5601C382006B CRC64;

Query Match 27.8%; Score 27; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 6.3e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLNSK 8
Db :||| |:::
4 QPGHLTSR 11

RESULT 90
Q15931 PRELIMINARY; PRT; 38 AA.
ID Q15931
AC Q15931;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNA-binding protein (Fragment).
GN Name=ZNF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92052132; PubMed=1946370;
RA Bray P., Lichter P., Thiesen H.J., Ward D.C., Dawid I.B.;
RT "Characterization and mapping of human genes encoding zinc finger
RT proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9563-9567(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92372070; PubMed=1505991;
RA Lichter P., Bray P., Ried T., Dawid I.B., Ward D.C.;
RT "Clustering of C2-H2 zinc finger motif sequences within telomeric and
RT fragile site regions of human chromosomes.";
RL Genomics 13:999-1007(1992).
DR EMBL; M88368; AAA61326.1; -
DR GO; GO:0003677; F:DNA binding; IEA.
KW DNA-binding.
FT NON_TER 1 1
FT NON_TER 38 38
SQ SEQUENCE 38 AA; 4284 MW; 001AE073A9DD9740 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 38;
Best Local Similarity 57.1%; Pred. No. 6.3e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLNS 7
Db :||| |:::
2 EKNHMA 8

RESULT 91
Q9GAU0 PRELIMINARY; PRT; 38 AA.
ID Q9GAU0
AC Q9GAU0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ATPase subunit 8 (Fragment).
OS Rana sylvatica (Wood frog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=45438;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Wu S.-B., Storey K.B.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBSJ databases.
CC -!- FUNCTION: This is one of the chains of the nonenzymatic component
CC (CF(0) subunit) of the mitochondrial ATPase complex (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).

DT	01-MAY-2000	(TREMELrel. 13, Created)
DT	01-MAY-2000	(TREMELrel. 13, Last sequence update)
DT	01-JUN-2002	(TREMELrel. 21, Last annotation update)
DE	BdrC2-	
GN	Name=bdrC2-;	
OS	Borrelia hermsli.	
OC	Bacteria; Spirochaetes; Spirochaetales; Borrelia.	
OX	NCBI_TaxID=140;	
RN	[1]	

RESULT	95
Q7QWJ3	
ID	Q7QWJ3
PRELIMINARY;	PRT; 42 AA.
AC	Q7QWJ3;
DT	01-MAR-2004 (TRENBLrel. 26, Created)
DT	01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT	01-MAR-2004 (TRENBLrel. 26, Last annotation update)

<hr/>					
RESULT	97				
Q7P9P5					
ID	Q7P9P5	PRELIMINARY;	PRT;	44 AA.	
AC	Q7P9P5;				
DT	01-MAR-2004	(TrEMBLrel. 26, Created)			
DT	01-MAR-2004	(TrEMBLrel. 26, Last sequence update)			
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)			
DE	Hypothetical protein.				
GN	Name=rsib_orf.963;				
OS	Rickettsia sibirica.				
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;				
OC	Rickettsiaceae; Rickettsiae; Rickettsia.				
OX	NCBI_TaxID=35793;				
[1]					
RP	SEQUENCE FROM N.A.				
RA	Malek J.A., Ereemeeva M.E., Daech G.A.;				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
CC	-!- CAUTION: The sequence shown here is derived from an				
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is				
CC	preliminary data.				
DR	EMBL; AABW01000001; EAA26145.1; -.				
KW	Hypothetical protein.				
SQ	SEQUENCE	44 AA;	5109 MW;	5CB6A274B57DC80B CRC64;	
<hr/>					
Query Match		27.8%;	Score 27;	DB 2;	Length 44;
Best Local Similarity		31.6%;	Pred. No. 7.3e+03;		
Matches		6;	Conservative	7;	Mismatches 4;
					Indels 2;
					Gaps
Qy	1 EPNHLNSKI--AFKIVSQE 17				
	: : : : : : :				
Dd	3 KPYVINEKLGAAYKANQD 21				
<hr/>					
RESULT	98				
Q72AJ8					
ID	Q72AJ8	PRELIMINARY;	PRT;	44 AA.	
AC	Q72AJ8;				
DT	05-JUL-2004	(TrEMBLrel. 27, Created)			
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)			
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)			
DE	Hypothetical protein.				
DE	Ordered locus Names=DVU11994;				
GN	Desulfotribrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB				
OS	8303).				
OC	Bacteria; Proteobacteria; Deltaproteobacteria; Desulfotribriales;				
OC	Desulfotribriaceae; Desulfotribriales;				
OX	NCBI_TaxID=882;				
[1]					
RP	SEQUENCE FROM N.A.				
RA	PubMed=15077118; DOI=10.1038/nbt959;				
RA	Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,				
RA	Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,				
RA	Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,				
RA	Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,				
RA	Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,				
RA	Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,				
RA	Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;				
RT	"The genome sequence of the anaerobic, sulfate-reducing bacterium				
RT	Desulfotribrio vulgaris Hildenborough."				
RL	Nat. Biotechnol. 22:554-559(2004).				
DR	EMBL; AE017315; AAS96470.1; -.				
DR	TIGR; DVU1994; -.				
KW	Complete proteome; Hypothetical protein.				
SQ	SEQUENCE	44 AA;	5059 MW;	BF626D0A31D46E3C CRC64;	
<hr/>					
Query Match		27.8%;	Score 27;	DB 2;	Length 44;
Best Local Similarity		45.5%;	Pred. No. 7.3e+03;		
Matches		5;	Conservative	2;	Mismatches 4;
					Indels 0;
					Gaps
Qy	1 EPNHLNSKIAP 11				
Dd	31 EERLNSRVHF 41				

```
RA Dutreix M., Backman A., Celerier J., Bagdasarian M.M., Sommer S.,
RA Bailone A., Devoret R., Bagdasarian M.;
RT "Identification of psiB genes of plasmids F and R6-5. Molecular basis
RT for psiB enhanced expression in plasmid R6-5.";
RL Nucleic Acids Res. 16:10869-10679(1988).
RN [4]
RN SEQUENCE FROM N.A.
RP MEDLINE=90136505; PubMed=2693941;
RX Loh S., Cram D., Skurray R.;
RA "Nucleotide sequence of the leading region adjacent to the origin of
RT transfer on plasmid F and its conservation among conjugative
RT plasmids.";
RL Mol. Gen. Genet. 219:177-186(1989).
RN [5]
RN SEQUENCE FROM N.A.
RP MEDLINE=90356405; PubMed=2201950;
RX Loh S., Skurray R., Celerier J., Bagdasarian M., Bailone A.,
RA Devoret R.;
RT "Nucleotide sequence of the psiA (plasmid SOS inhibition) gene located
RT on the leading region of plasmids F and R6-5.";
RL Nucleic Acids Res. 18:4597-4597(1990).
RN [6]
RN SEQUENCE FROM N.A.
RP MEDLINE=9296678; PubMed=10366527;
RX Manwaring N.P., Skurray R.A., Firth N.;
RA "Nucleotide sequence of the F plasmid leading region.";
RT Plasmid 41:219-225(1999).
RL EMBL; AF106329; A047190.1; -.
DR Hypothetical protein; Plasmid.
KW SEQUENCE 45 AA; 5211 MW; 8F8AC32FFADBE85 CRC64;
SQ
Query Match 27.8%; Score 27; DB 2; Length 45;
Best Local Similarity 40.0%; Pred. No. 7.5e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 3 NLSNLSKIAPK 12
DB 11 NLSNLSRVHVF 20
RESULT 101
Q83CF9 PRELIMINARY; PRT; 45 AA.
AC Q83CF9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=CHU1159;
OS Coccidia burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coccidiaceae; Coccidia.
OX NCBI_TaxID=777;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Nine Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=12704232;
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin H., Davidson T.M., Beanan M.J.,
RA DeBoy R.T., Dougherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coccidia
RT burnetii.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR EMBL; AE016963; AA090670.1; -.
TX TIGR; CHU1159; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 45 AA; 5331 MW; 9BDD801CBC5F6362 CRC64;
Query Match 27.8%; Score 27; DB 2; Length 45;
Best Local Similarity 62.5%; Pred. No. 7.5e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
RA AAS96470 PRELIMINARY; PRT; 44 AA.
AC AAS96470;
DT 27-APR-2004 (TrEMBLrel. 27, Created)
DT 27-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN DVU1994.
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RN SEQUENCE FROM N.A.
RP PubMed=15077118;
RX Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N., Methe B.A., Brinkac L.M.,
RA Dougherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Uitterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017315; AAS96470.1; -.
TX TIGR; DVU1994; -.
KW Hypothetical protein.
SQ SEQUENCE 44 AA; 5059 MW; BF626D0A31D46B3C CRC64;
Query Match 27.8%; Score 27; DB 2; Length 44;
Best Local Similarity 45.5%; Pred. No. 7.3e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 EPNHLSKIAP 11
DB 31 EERNLSRVHVF 41
RESULT 100
Q9S4W1 PRELIMINARY; PRT; 45 AA.
AC Q9S4W1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Escherichia coli.
OG Plasmid F.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=84236116; PubMed=6329741;
RX Thompson R., Taylor L., Kelly K., Everett R., Willetts N.;
RA "The F plasmid origin of transfer: DNA sequence of wild-type and
RT mutant origins and location of origin-specific nicks.";
RL EMBO J. 3:1175-1180(1984).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=89006267; PubMed=3049248;
RX Loh S.M., Cram D.S., Skurray R.A.;
RA "Nucleotide sequence and transcriptional analysis of a third function
RT (Flm) involved in F-plasmid maintenance.";
RL Gene 66:259-268(1988).
RN [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=89083490; PubMed=3205720;
```

```

QY 1 EPNHLSK 8
| | | | |
Db 16 ESNHLHRK 23

RESULT 102
Q82Z68
ID Q82Z68 PRELIMINARY; PRT; 46 AA.
AC Q82Z68;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=EF3205;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=2250857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Banerjee L., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayar L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis,"
RL Science 299:2071-2074(2003).
DR EMBL; AE016957; AAC82877.1; -.
DR TIGR; EF3205; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 46 AA; 5557 MW; 4C4DA4FDEBD78FB5 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 46;
Best Local Similarity 40.8%; Pred. No. 7.6e+03;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 PNHLSKIAFKIVSQ 16
| | | | |
Db 22 PNITKTRKALKIIXE 36

RESULT 103
Q9BX55
ID Q9BX55 PRELIMINARY; PRT; 47 AA.
AC Q9BX55;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mitotic spindle checkpoint kinase (Fragment).
GN Name=BUB1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20424176; PubMed=10969775;
RA Jaffrey R.G., Pritchard S.C., Clark C., Murray G.I., Cassidy J.,
RA Kerr K.M., Nicolson M.C., McLeod H.L.;
RT "Genomic instability at the BUB1 locus in colorectal cancer, but not
RT in non-small cell lung cancer,"
RL Cancer Res. 60:4349-4352(2000).
DR EMBL; AF264055; AAK29549.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
KW Kinase.
FT NON_TER 1
FT NON_TER 47
SQ SEQUENCE 47 AA; 5425 MW; 79B68620617C1440 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 47;
Best Local Similarity 62.5%; Pred. No. 7.8e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSK 8
| | | | |
Db 30 EPNKNTK 37

RESULT 105
Q8FDQ3
ID Q8FDQ3 PRELIMINARY; PRT; 47 AA.
AC Q8FDQ3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein c3685.
GN OrderedLocusNames=c3685;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=O6:HL / CFT073 / ATCC 700928;
RX MEDLINE=22386234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.J., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic *Escherichia coli*.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016766; AAN82133.1; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 47 AA; 5344 MW; FA35BE0184673DAB CRC64;

Query Match 27.8%; Score 27; DB 2; Length 47;
Best Local Similarity 54.5%; Pred. No. 7.8e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVS 15
||| :||
Db 36 LNRKITWVILS 46
||| :||

RESULT 106
Q9X1D2 PRELIMINARY; PRT; 47 AA.
AC Q9X1D2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=TM1412;
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=92287316; PubMed=10360571; DOI=10.1038/20601;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L.A., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,
RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of the thermotoga maritima.";
RL Nature 393:323-329 (1999).
DR EMBL; AE001793; AAD36483.1; -;
DR PIR; E72258; E72258.
DR TIGR; TM1412; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 47 AA; 5438 MW; 776C75B7B97866BA CRC64;

Query Match 27.8%; Score 27; DB 2; Length 47;
Best Local Similarity 38.5%; Pred. No. 7.8e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKI 13
||| :||
Db 2 ETHLLNSVVCYKL 14
||| :||

RESULT 107
Q6L7B1 PRELIMINARY; PRT; 47 AA.
AC Q6L7B1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Beta2-microglobulin (Fragment).
GN Name=B2m;
OS Ctenopharyngodon idella (Grass carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Ctenopharyngodon.
OX NCBI_TaxID=7959;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Yang T.Y., Xia C.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RE EMBL; AB128863; BAD22757.1; -;
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 47 AA; 5594 MW; B8C8A85E440C9054 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 47;
Best Local Similarity 44.4%; Pred. No. 7.8e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HLNSKIAPK 12
||| :||
Db 13 HLTKSVSFK 21
||| :||

RESULT 108
ID_PSD AZOVI STANDARD; PRT; 48 AA.
AC Q44558;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Phosphatidylserine decarboxylase proenzyme (EC 4.1.1.65) [Contains:
DE Phosphatidylserine decarboxylase beta chain] (Fragment).
GN Name=psd;
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=OP / UW136;
RX MEDLINE=96184904; PubMed=8617271;
RA Collnaghi R., Pagani S., Kennedy C., Drummond M.;
RA "Cloning, sequence analysis and overexpression of the rhodanese gene
of Azotobacter vinelandii.";
RL Eur. J. Biochem. 236:240-248 (1996).
CC -!- CATALYTIC ACTIVITY: Phosphatidyl-L-serine =
phosphatidylethanolamine + CO(2).
CC -!- COFACTOR: Pyruvoyl group (By similarity).
CC -!- SIMILARITY: Belongs to the phosphatidylserine decarboxylase
family. Subfamily 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; L42346; AAB03238.1; -;
DR PIR; S62189; S62189.
DR HAMAP; MF_00662; -; 1.
KW Decarboxylase; Lyase; Phospholipid biosynthesis; Pyruvate.
FT CHAIN 1 >48 Phosphatidylserine decarboxylase beta
chain (By similarity).
FT NON TER 48
SQ SEQUENCE 48 AA; 5591 MW; 366CC6E30EC3F699 CRC64;

Query Match 27.8%; Score 27; DB 1; Length 48;
Best Local Similarity 41.7%; Pred. No. 8e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PNLHNSKIAFKI 13
 ID : : : : :
 Db 14 PHLHLSRAAGRL 25

RESULT 109

Q6PVS2 PRELIMINARY; PRT; 48 AA.
 AC Q6PVS2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Interleukin 5 (Fragment).
 GN Name=IL-5;
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Abomasal biopsy;
 RA Shanmuganathan T., Hickford J.G.H., Savill M., Sykes A.R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY575612; AAS87353.1; -.
 DR InterPro; IPR009079; 4 helix cytokine.
 DR PRINTS; PR00432; INTERLEUKIN5.
 DR ProDom; PD006721; Interleukin_5; 1.
 FT NON_TER 1 1
 FT NON_TER 48 48
 SQ SEQUENCE 48 AA; 5256 MW; 99486B0E54779B24 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 48;
 Best Local Similarity 33.3%; Pred. No. 8e+03;
 Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PNLHNSKIAFKIVSQ 16
 ID : : : : :
 Db 20 PQHTNHCLEEVFQ 34

RESULT 110

Q8KE64 PRELIMINARY; PRT; 48 AA.
 AC Q8KE64;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein CT0826.
 GN OrderedLocusNames=CT0826;
 OS Chlorobium tepidum.
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
 OC Chlorobaculum.
 OC NCBI_TaxID=1097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TLS / ATCC 49652 / DSM 12025;
 RC MEDLINE=22103695; PubMed=12093901; DOI=10.1073/pnas.132181499;
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
 RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
 RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
 RA Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
 RA Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,
 RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
 RT "The complete genome sequence of Chlorobium tepidum TLS, a
 RT photosynthetic, anaerobic, green-sulfur bacterium."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
 DR EMBL; AE012850; AAM72062.1; -.
 DR TIGR; CT0826; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 48 AA; 5675 MW; 68C0BFED89EC8C3 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 48;
 Best Local Similarity 35.3%; Pred. No. 8e+03;
 Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 3 PNLHNSKIAFKIVSQEPA 19
 ID : : : : :
 Db 13 NRIDRKIAMALAFRPA 29

RESULT 111

AAS87353 PRELIMINARY; PRT; 48 AA.
 AC AAS87353;
 DT 15-APR-2004 (TrEMBLrel. 27, Created)
 DT 15-APR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 15-APR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Interleukin 5 (Fragment).
 GN IL-5.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Abomasal biopsy;
 RA Shanmuganathan T., Hickford J.G.H., Savill M., Sykes A.R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY575612; AAS87353.1; -.
 FT NON_TER 1 1
 FT NON_TER 48 48
 SQ SEQUENCE 48 AA; 5256 MW; 99486B0E54779B24 CRC64;

Query Match 27.8%; Score 27; DB 2; Length 48;
 Best Local Similarity 33.3%; Pred. No. 8e+03;
 Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PNLHNSKIAFKIVSQ 16
 ID : : : : :
 Db 20 PQHTNHCLEEVFQ 34

RESULT 112

Q9AGM9 PRELIMINARY; PRT; 50 AA.
 AC Q9AGM9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE GMP459 (Fragment).
 GN Name=gmp459;
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON=CW459tet;
 RC MEDLINE=21220755; PubMed=11320127;
 RA Roberts A.P., Johansen P.A., Lytras D., Mullany P., Rood J.I.;
 RT "Comparison of Tn3397 from Clostridium difficile, Tn916 from
 RT Enterococcus faecalis and the CW459tet(M) element from Clostridium
 RT perfringens shows that they have similar conjugation regions but
 RT different insertion and excision modules."
 RL Microbiology 147:1243-1251(2001).
 DR EMBL; AF323848; AAK17959.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003922; F:GMP synthase (glutamine-hydrolyzing) activity; IEA.
 DR GO; GO:0006177; P:GMP biosynthesis; IEA.
 DR GO; GO:0006164; P:purine nucleotide biosynthesis; IEA.
 DR InterPro; IPR001674; GMP synth_C.
 DR Pfam; PF00958; GMP_synt_C; 1.

```

FT NON_TER 1 1
SQ SEQUENCE 50 AA; 5689 MW; 58A68ACE792DCAAC CRC64;

Query Match 27.8%; Score 27; DB 2; Length 50;
Best Local Similarity 30.8%; Pred. No. 8.3e+03;
Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SKIAFKIVSQEPA 19
   :| :| :| :| :| :|
Db 33 NRIVYDVTSKPPA 45

RESULT 113
Q7VLJ7 PRELIMINARY; PRT; 32 AA.
AC Q7VLJ7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS OrderedLocusNames=HD1434;
OC Bacteriophage;
OC Haemophilus ducreyi.
OC Pasteurellales;
OC Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RA "The complete genome sequence of Haemophilus ducreyi";
RA Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB017154; AAF96240.1; -.
RW Complete proteome; Hypothetical protein.
SQ SEQUENCE 32 AA; 3832 MW; 57E579D506BEA82 CRC64;

Query Match 27.3%; Score 26.5; DB 2; Length 32;
Best Local Similarity 42.9%; Pred. No. 6.4e+03;
Matches 6; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy 3 NNLNSKI-AFKIVS 15
   || :| :| :| :| :|
Db 12 NHVKKQVFAFIIS 25

RESULT 114
Q9RHE3 PRELIMINARY; PRT; 36 AA.
AC Q9RHE3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mbc-like protein.
OS Pedicoccus pentosaceus.
OG Pedicoccus pentosaceus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae; Pedicoccus.
OX NCBI_TaxID=1255;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=ATCC43200;
RX MEDLINE=20153810; PubMed=10686129;
RA Giacomini A., Squartini A., Nuti M.P.;
RA "Nucleotide sequence and analysis of plasmid pMD136 from Pedicoccus
RT pentosaceus FBE61 (ATCC43200) involved in pediocin A production.";
RL Plasmid 43:111-122 (2000).
RN [2]
RC SEQUENCE FROM N.A.
RA STRAIN=ATCC43200;
RA Giacomini A., Marazzan G., Salvato P., Squartini A., Nuti M.P.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF033858; AAF22855.1; -.
RW Plasmid.
SQ SEQUENCE 36 AA; 4372 MW; ACFB9CC1D9993C87 CRC64;

```

```

Query Match 27.3%; Score 26.5; DB 2; Length 36;
Best Local Similarity 33.3%; Pred. No. 7.2e+03;
Matches 6; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

Qy 1 EPNHLSK-IAFKIVSQE 17
   :| :| :| :| :| :| :|
Db 15 QPNKESQINFRVSEQD 32

RESULT 115
Q9PCI9 PRELIMINARY; PRT; 37 AA.
AC Q9PCI9
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Xf1790;
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=9A5C;
RA MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lenos E.G.M., Lenas M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C. de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159 (2000).
DR EMBL; AE004001; AAF84598.1; -.
DR PIR; D82636; D82636.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 37 AA; 3952 MW; D2656E6BA48B9E CRC64;

Query Match 27.3%; Score 26.5; DB 2; Length 37;
Best Local Similarity 41.2%; Pred. No. 7.4e+03;
Matches 7; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy 2 EPNHLSKIAFKIVSQEP 18
   || :| :| :| :| :| :|
Db 15 FNLVNGQIAC-VLGGSP 30

RESULT 116
Q727X8 PRELIMINARY; PRT; 39 AA.
ID Q727X8
AC Q727X8;

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=DVU2726;
 OS Desulfobrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobirionales;
 OC Desulfobirionaceae; Desulfobrio.
 OC NCBI_TaxID=882;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15077118; DOI=10.1038/nbt959;
 RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
 RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
 RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
 RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
 RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
 RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 Desulfobrio vulgaris Hildenborough.";
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL; AB017318; AAS97198.1; -;
 DR TIGR; DVU2726; -;
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 39 AA; 4271 MW; 5002AE73D554CF1D CRC64;

 Query Match 27.3%; Score 26.5; DB 2; Length 39;
 Best Local Similarity 39.9%; Pred. No. 7.8e+03;
 Matches 7; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

 QY 2 PNHLNSK-IAPKIVSQEP 18
 Db 19 PFHLSSPSATRVAAWEP 36

 RESULT 117
 AAS97198
 ID AAS97198 PRELIMINARY; PRT; 39 AA.
 AC AAS97198
 DT 26-APR-2004 (TrEMBLrel. 27, Created)
 DT 26-APR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 11-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN DVU2726.
 OS Desulfobrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobirionales;
 OC Desulfobirionaceae; Desulfobrio.
 OC NCBI_TaxID=882;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15077118;
 RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
 RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
 RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
 RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
 RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
 RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 Desulfobrio vulgaris Hildenborough.";
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL; AB017318; AAS97198.1; -;
 DR TIGR; DVU2726; -;
 KW Hypothetical protein.
 SQ SEQUENCE 39 AA; 4271 MW; 5002AE73D554CF1D CRC64;

 Query Match 27.3%; Score 26.5; DB 2; Length 39;
 Best Local Similarity 38.9%; Pred. No. 7.8e+03;
 Matches 7; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 2 PNHLNSK-IAPKIVSQEP 18
 Db 19 PFHLSSPSATRVAAWEP 36

 RESULT 118
 Q99J74
 ID Q99J74 PRELIMINARY; PRT; 41 AA.
 AC Q99J74
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Karyaria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallaloo D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005705; AAF05705.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 41 AA; 4720 MW; 7692D458655D3B05 CRC64;

 Query Match 27.3%; Score 26.5; DB 2; Length 41;
 Best Local Similarity 43.8%; Pred. No. 8.2e+03;
 Matches 7; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

 QY 3 NHLNSKIAFKIVSQEP 18
 Db 27 NH-GLKVARSLGSRP 41

 RESULT 119
 Q9BX20
 ID Q9BX20 PRELIMINARY; PRT; 19 AA.
 AC Q9BX20
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE BA470C13.1 (KIAA1272 similar to rat tulip proteins 1 and 2)
 DE (Fragment).
 GN Name=BA470C13.1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RA  Peck A.;
RL  Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AL161658; CAC36066.1; -.
FT  NON-TER 1
SQ  SEQUENCE 19 AA; 2211 MW; 2BB9DB0ABFF5A5C0 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 19;
Best Local Similarity 36.4%; Pred. No. 4.5e+03;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 8 KIAPKIVSQEP 18
Db 5 KVSFQYLSKGP 15

RESULT 120
Q9X3G5 PRELIMINARY; PRT; 21 AA.
AC Q9X3G5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Cytochrome b6/f complex subunit IV (Fragment).
GN Name=petD;
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630 (1998).
DR EMBL; AF070159; AAD23205.1; -.
FT NON-TER 21
SQ SEQUENCE 21 AA; 2283 MW; CC97B526F03C474F CRC64;

Query Match 26.8%; Score 26; DB 2; Length 21;
Best Local Similarity 38.5%; Pred. No. 5e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKI 13
Db 6 KPNLADSKIRAKL 18

RESULT 121
Q9LA84 PRELIMINARY; PRT; 23 AA.
AC Q9LA84;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RA Cross R., Aish J., Paston S.J., Poole R.K., Moir J.W.;
RT "Cytochrome c' from Rhodobacter capsulatus confers increased resistance to nitric oxide.";
RL J. Bacteriol. 182:1442-1447 (2000).
DR EMBL; AF147705; AAF37604.1; -.
FT NON-TER 1
SQ SEQUENCE 23 AA; 2271 MW; FEE78307DEDDFA00 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 23;
Best Local Similarity 27.8%; Pred. No. 5.5e+03;
Matches 5; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 2 PNHLSKIAPKIVSQEPA 19
Db 5 PGHSGKAALIDLVAVKPA 22

RESULT 122
YCXA ODOSI STANDARD; PRT; 27 AA.
AC P49836;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 3.3 kDa protein in rpl11-trnW intergenic region (ORF27).
OS Odontella sinensis (Marine centric diatom).
OG Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscaceae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga, Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342 (1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z67753; CAA91728.1; -.
DR PIR; S78355; S78355.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 27 AA; 3297 MW; 762B05DA25B8D2D3 CRC64;

Query Match 26.8%; Score 26; DB 1; Length 27;
Best Local Similarity 46.7%; Pred. No. 6.5e+03;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 5 LNSKIAPKIVSQEPA 19
Db 8 LRKKIAEVWIFQNEA 22

RESULT 123
Q9ZG21 PRELIMINARY; PRT; 29 AA.
AC Q9ZG21;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Heat shock protein (fragment).
GN Name=htrA;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF087355; AAD04127.1; -.
DR EMBL; AF087355; AAD04127.1; -.
DR GO; GO:0003773; F:heat shock protein activity; IEA.
DR InterPro; IPR009003; Pept_Ser_Cys.
KW Heat shock.
```



```

FT NON_TER 1 1
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3188 MW; E997FCA4F95B9A5D CRC64;

Query Match 26.8%; Score 25; DB 2; Length 29;
Best Local Similarity 40.0%; Pred. No. 7e+03;
Matches 6; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 4 HLNKSKIAFKIVSQEP 18
Db 6 HDGKKYAKIVGLDP 20

RESULT 124
Q7SOM7
ID Q7SOM7 PRELIMINARY; PRT; 31 AA.
AC Q7SOM7;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE Predicted protein.
GN Name=NCU05913.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysseis M., Mauceli E., Bielek C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Sailer S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.
RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000489; EAA28877.1; -.
SQ SEQUENCE 31 AA; 3649 MW; A22BF1062BFD0E58 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 31;
Best Local Similarity 23.5%; Pred. No. 7.5e+03;
Matches 4; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NHLNLSKIAFKIVSQEPA 19
Db 11 HHMSAKAFNRELGREPS 27

RESULT 125
Q9XSB9
ID Q9XSB9 PRELIMINARY; PRT; 31 AA.
AC Q9XSB9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE Thymidine DNA glycosylase (Fragment).
OS Ateles belzebuth chamek (Chamek spider monkey).
OC Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.
OX NCBI_TaxID=118643;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20169179; PubMed=10702662;
RA Canavez F.C., Moreira M.A., Bonvicino C.R., Parham P., Seuanez H.N.;
RA "Evolutionary disruptions of human syntenic groups 3, 12, 14, and 15
RA in Ateles paniscus chamek (Platyrrhini, primates).";
RL Cytogenet. Cell Genet. 87:182-188(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Moreira M.A.M.;
RA Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF117602; AAD25362.1; -.
DR InterPro; IPR008967; P53_like_DNA_bnd.
FT NON_TER 1 1
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 3698 MW; E139811C3F581519 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 31;
Best Local Similarity 41.7%; Pred. No. 7.5e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 6 NSKIAFKIVSQE 17
Db 17 NGKCIYEIFSKE 28

RESULT 126
Q9KM48
ID Q9KM48 PRELIMINARY; PRT; 31 AA.
AC Q9KM48;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein VCA0541.
GN OrderedLocuNames=VCA0541;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RC MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RA cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AEC04385; AAF96443.1; -.
DR PIR; E82446; E82446.
DR TIGR; VCA0541; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 31 AA; 3688 MW; 55FA0116907D6D2D CRC64;

Query Match 26.8%; Score 26; DB 2; Length 31;
Best Local Similarity 80.0%; Pred. No. 7.5e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NHLNS 7
Db 6 NHINS 10

RESULT 127
Q8E8Y3
ID Q8E8Y3 PRELIMINARY; PRT; 34 AA.
AC Q8E8Y3;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

```

RESULT 129					
ID	Q7U051	PRELIMINARY;	PRT;	35 AA.	
AC	Q7U051;				
DT	01-OCT-2003 (TrEMBLrel. 25, Created)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)				
DE	Hypothetical protein.				
GN	OrderedLocusNames=RB3505;				
OS	Rhodopirellula baltica.				
OC	Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;				
OC	Planctomycetaceae; Pirellula.				
OX	NCBI_TaxID=117;				
OX	[1]				
RN	SEQUENCE FROM N.A.				
RP	STRAIN=1;				
RC	MEDLINE=22735913; PubMed=12835416;				
RX	Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,				
RA	Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,				
RA	Schlesner H., Amann R., Reinhardt R.;				
RT	"Complete genome sequence of the marine planctomycete Pirellula sp strain 1.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003) .				
DR	EMBL; BX294138; CAD73233.1; -.				
DR	InterPro; IPR011479; DUF1586.				
DR	Pfam; PF07625; DUF1586; 1.				
KW	Complete proteome; Hypothetical protein.				
QY	SEQUENCE 35 AA; 3664 MW; 74400E8A535E7455 CRC64;				
Query Match	26.8%;	Score 26;	DB 2;	Length 35;	
Best Local Similarity	50.0%;	Pred. No. 8.5e+03;			
Matches	6; Conservative	1; Mismatches	5; Indels	0; Gap	
Qy	7 SKTAFKIVGQEP 18 : 3 SRTALAAVQSQT 14				
RESULT 130					
ID	Q7UZ39	PRELIMINARY;	PRT;	35 AA.	
AC	Q7UZ39;				
DT	01-OCT-2003 (TrEMBLrel. 25, Created)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)				
DE	Hypothetical protein.				
GN	OrderedLocusNames=RB206;				
OS	Rhodopirellula baltica.				
OC	Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;				
OC	Planctomycetaceae; Pirellula.				
OX	NCBI_TaxID=117;				
OX	[1]				
RN	SEQUENCE FROM N.A.				
RP	STRAIN=1;				
RC	MEDLINE=22735913; PubMed=12835416;				
RA	Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,				
RA	Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,				
RA	Schlesner H., Amann R., Reinhardt R.;				
RT	"Complete genome sequence of the marine planctomycete Pirellula sp strain 1.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003) .				
DR	EMBL; BX294133; CAD71447.1; -.				
DR	InterPro; IPR011479; DUF1586.				
DR	Pfam; PF07625; DUF1586; 1.				
KW	Complete proteome; Hypothetical protein.				
SO	SEQUENCE 35 AA; 3818 MW; 71C14AF152EBB130 CRC64;				

```

QY 7 SKIAFKIVSQP 18
Db 3 SRTALAAVSQTP 14

RESULT 131
Q8EXS3 PRELIMINARY; PRT; 35 AA.
AC Q8EXS3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LB135;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AE011602; AAN51694.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 35 AA; 3956 MW; CBF48BA44BCB190 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 35;
Best Local Similarity 50.0%; Pred. No. 8.5e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NNLNSKIAFK 12
Db 26 NLNSKVSXK 35

RESULT 132
Q9KNU1 PRELIMINARY; PRT; 35 AA.
AC Q9KNU1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VC2639.
GN OrderedLocusNames=VC2639;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004330; AAF95780.1; -.
FIR; F82051; F82051.

RESULT 133
Q7M2T4 PRELIMINARY; PRT; 38 AA.
AC Q7M2T4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Properdin (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=86225592; PubMed=3635564;
RA Nakano Y., Matsuda T., Sakamoto T., Tomita M.;
RT "Isolation and characterization of rabbit properdin of the alternative
RT complement pathway.";
RL J. Immunol. Methods 90:77-83(1986).
DR PIR; A60867; A60867.1
FT NON_TER 1
FT NON_TER 38
SQ SEQUENCE 38 AA; 4031 MW; 20CE2B37661D1A42 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 9.3e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFK 12
Db 27 EPCLNAYAFQ 38

RESULT 134
Q6J5M5 PRELIMINARY; PRT; 38 AA.
AC Q6J5M5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98-D60C;
RA Dmitriev A., Shen A.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY598359; AAT10375.1; -.
KW Hypothetical protein.
SQ SEQUENCE 38 AA; 4579 MW; 3B1715659362A642 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 40.0%; Pred. No. 9.3e+03;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 PNHLNSKIAFKIVSQ 16

```



```

SQ SEQUENCE 38 AA; 4579 MW; 3B1715659362A642 CRC64;
Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 40.0%; Pred. No. 9.3e+03;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PNHLNSKIAFKIVSQ 16
||| | | | | |
17 PNIXTRKALKIIE 31

Db

RESULT 139
Q6EZK5 PRELIMINARY; PRT; 38 AA.
ID Q6EZK5
AC Q6EZK5;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=GBAA_PX01_0041;
OS Bacillus anthracis.
OG Plasmid PX01.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017336; AAT28782.2; -
DR TIGR; GBAA_PX01_0041; -
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 38 AA; 4624 MW; 61179E72D8D684F4 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 33.3%; Pred. No. 9.3e+03;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQ 16
:|:|:|:|:|
1 MHSKISYKYTNR 12

Db

RESULT 140
Q6VYR1 PRELIMINARY; PRT; 38 AA.
ID Q6VYR1
AC Q6VYR1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Heart-type fatty acid binding protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ye M., Cao H., Wen J.;
RT "RFLPs at Heart and Adipocyte Fatty Acid Binding Protein Genes in
RT Beijing Oil Chicken.";
RL Anim. Biotechnol. Bull. 8:88-94 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Ye M., Cao H., Wen J.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY319283; AAQ24606.1; -
DR EMBL; AY319284; AAQ24607.1; -
DR EMBL; AY319281; AAQ24604.1; -
InterPro; IPR011038; Calycin.

```

```

FT NON_TER 1 1
FT NON_TER 38 38
SQ SEQUENCE 38 AA; 4314 MW; 6F45B7C06DB7E204 CRC64;
Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 9.3e+03;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 NSKIAFKI 13
|:|:|:|:|
5 NTEISFKL 12

Db

RESULT 141
AAQ24604 PRELIMINARY; PRT; 38 AA.
ID AAQ24604
AC AAQ24604;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Heart-type fatty acid binding protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ye M., Cao H., Wen J.;
RT "RFLPs at Heart and Adipocyte Fatty Acid Binding Protein Genes in
RT Beijing Oil Chicken.";
RL Anim. Biotechnol. Bull. 8:88-94 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Ye M., Cao H., Wen J.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY319281; AAQ24604.1; -
FT NON_TER 1 1
FT NON_TER 38 38
SQ SEQUENCE 38 AA; 4314 MW; 6F45B7C06DB7E204 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 9.3e+03;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 NSKIAFKI 13
|:|:|:|:|
5 NTEISFKL 12

Db

RESULT 142
AAQ24606 PRELIMINARY; PRT; 38 AA.
ID AAQ24606
AC AAQ24606;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Heart-type fatty acid binding protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ye M., Cao H., Wen J.;
RT "RFLPs at Heart and Adipocyte Fatty Acid Binding Protein Genes in
RT Beijing Oil Chicken.";
RL Anim. Biotechnol. Bull. 8:88-94 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Ye M., Cao H., Wen J.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

```

Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 40.0%; Pred. No. 9.3e+03;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Query 2 PNHLNSKIAFKIVSQ 16
DB 17 PNKTRKRAALKIIEH 31

RESULT 145
AAT28782 PRELIMINARY; PRT; 38 AA.
AC AAT28782;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN GBAA_PX01_0041.
OS Bacillus anthracis str. Ames 0581.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus anthracis.
OX NCBI_TaxID=261594;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S., Fraser C.M.;
RT "Bacillus anthracis comparative genomics."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF019336; AAT28782.2; --
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 38 AA; 4624 MW; 61179E72D8D684F4 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 33.3%; Pred. No. 9.3e+03;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Query 5 LNSKIAFKIVSQ 16
DB 1 MESHISYKYTNR 12

RESULT 146
Q8N4V7 PRELIMINARY; PRT; 39 AA.
AC Q8N4V7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 9.3e+03;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Query 6 NSKIAFKI 13
DB 5 NTEISFKL 12

RESULT 143
AAQ24607 PRELIMINARY; PRT; 38 AA.
ID AAQ24607;
AC AAQ24607;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Heart-type fatty acid binding protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ye M., Cao H., Wen J.;
RT "RFLPs at Heart and Adipocyte Fatty Acid Binding Protein Genes in
RT Beijing Oil Chicken."
RL Anim. Biotechnol. Bull. 8:88-94 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Ye M., Cao H., Wen J.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY319284; AAQ24607.1; --
FT NON_TER 1 1
FT NON_TER 38 38
SQ SEQUENCE 38 AA; 4314 MW; 6F45B7C06DB7E204 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 9.3e+03;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Query 6 NSKIAFKI 13
DB 5 NTEISFKL 12

RESULT 144
AAT10375 PRELIMINARY; PRT; 38 AA.
ID AAT10375;
AC AAT10375;
DT 20-MAY-2004 (TrEMBLrel. 27, Created)
DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98-D60C;
RA Dmitriev A., Shen A.;
RT "Acquisition of the beta-antigen gene by Streptococcus agalactiae."
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY598359; AAT10375.1; --
KW Hypothetical protein.
SQ SEQUENCE 38 AA; 4579 MW; 3B1715659362A642 CRC64;

```

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033310; AAH33310.1; -.
KW Hypothetical protein.
SQ SEQUENCE 39 AA; 4649 MW; E2866DEC92884381 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 39;
Best Local Similarity 50.0%; Pred. No. 9.5e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 HLNKIAFKI 13
Db :|||||:
12 NVNSKRAFTV 21

RESULT 147
Q7VHY8 PRELIMINARY; PRT; 39 AA.
AC Q7VHY8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=HH0823;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droge M., Farmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906 (2003).
DR EMBL; AE017146; AAP77420.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 39 AA; 4704 MW; 14D4046FCCC74BAA CRC64;

Query Match 26.8%; Score 26; DB 2; Length 39;
Best Local Similarity 50.0%; Pred. No. 9.5e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 NNLKIAFKIV 14
Db :|||||:
25 NNLKILFILL 36

RESULT 148
Q8FZG8 PRELIMINARY; PRT; 39 AA.
AC Q8FZG8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BR1515;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.

```

```

OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
DR EMBL; AE014446; AAN30425.1; -.
DR TIGR; BR1515; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 39 AA; 4582 MW; 5FE3209B436C21C0 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 39;
Best Local Similarity 44.4%; Pred. No. 9.5e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 11 FKIVSQEPA 19
Db :|:|:|:|
3 FRILRRPA 11

RESULT 149
Q60573 PRELIMINARY; PRT; 39 AA.
AC Q60573;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Alpha-1-acid glycoprotein (Fragment).
GN Name=AGP;
OS Mus caroli (Wild mouse) (Ricefield mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10089;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90285134; PubMed=2354997;
RA Prowse K.R., Baumann H.;
RT "Molecular characterization and acute-phase expression of the multiple
RT Mus caroli alpha-1-acid glycoprotein (AGP) genes: Differences in
RT glucocorticoid stimulation and regulatory elements between the rat and
RT mouse AGP genes.";
RL J. Biol. Chem. 265:10201-10209 (1990).
DR EMBL; M34644; AAA37200.1; -.
FT NON TER 39
SQ SEQUENCE 39 AA; 4353 MW; 56229756A80C5715 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 39;
Best Local Similarity 33.3%; Pred. No. 9.5e+03;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PNHLNSKIAPKIVSQ 16
Db :|:|:|:|:|
21 PEHVNTIGEPITNE 35

RESULT 150
Q6RV10 PRELIMINARY; PRT; 40 AA.
AC Q6RV10;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative chimeric protein 173-56.
OS Trypanosoma cruzi.

```

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Teixeira A.R.L., Nitz N., Gomes C.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY490894; AAR88804.1; -
 SQ SEQUENCE 40 AA; 4736 MW; A4605A686F4B9043 CRC64;
 Query Match 26.8%; Score 26; DB 2; Length 40;
 Best Local Similarity 35.3%; Pred. No. 9.8e+03;
 Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
 Qy 2 PNHLNSKIAPKIVSQEP 18
 Db 18 PQFLNLRHLKLYPLP 34
 RESULT 151
 Q87077 PRELIMINARY; PRT; 40 AA.
 AC Q87077;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE Replicase motif region (ORF1b) (Fragment).
 OS Simian hemorrhagic fever virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Arteriviridae; Arterivirus.
 OX NCBI_TaxID=38143;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96435012; PubMed=8837898;
 RA Chen Z., Plagemann P.G.;
 RT "Detection of related positive-strand RNA virus genomes by reverse
 RT transcription/polymerase chain reaction using degenerate primers for
 RT common replicase sequences";
 RL Virus Res. 39:365-375(1995).
 DR EMBL; U28864; AAB40002.1; -
 FT NON TER 1 1
 FT NON TER 40 40
 SQ SEQUENCE 40 AA; 4450 MW; E76CBA2A26538928 CRC64;
 Query Match 26.8%; Score 26; DB 2; Length 40;
 Best Local Similarity 31.2%; Pred. No. 9.8e+03;
 Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 Qy 4 HLNSKIAPKIVSQEPA 19
 Db 4 HFATKLFEMACAEA 19
 RESULT 152
 Q91FA4 PRELIMINARY; PRT; 40 AA.
 ID Q91FA4
 AC Q91FA4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE 421L.
 OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
 OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
 OX NCBI_TaxID=10488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99125223; PubMed=9926400;
 RA Muller K., Tidona C.A., Bahr U., Darai G.;
 RT "Identification of a thymidylate synthase gene within the genome of
 RT Chilo iridescent virus";
 RL Virus Genes 17:243-258(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93118242; PubMed=1475907;

RA Sonntag K.C., Darai G.;
 RT "Characterization of the third origin of DNA replication of the genome
 RT of insect iridescent virus type 6.";
 RL Virus Genes 6:333-342(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94353641; PubMed=8073636;
 RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;
 RT "Chilo iridescent virus encodes a putative helicase belonging to a
 RT distinct family within the 'DEAD/H' superfamily: implications for the
 RT evolution of large DNA viruses";
 RL Virus Genes 8:151-158(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94292906; PubMed=8021587;
 RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
 RA Koonin E.V., Darai G.;
 RT "Insect iridescent virus type 6 encodes a polypeptide related to the
 RT largest subunit of eukaryotic RNA polymerase II.";
 RL J. Gen. Virol. 75:1557-1567(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95213160; PubMed=7698884;
 RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;
 RT "Identification of the primary structure and the coding capacity of
 RT the genome of insect iridescent virus type 6 between the genome
 RT coordinates 0.310 and 0.347 (7990 bp).";
 RL Intervirology 37:287-297(1994).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98141693; PubMed=9482589;
 RA Bahr U., Tidona C.A., Darai G.;
 RT "The DNA sequence of Chilo iridescent virus between the genome
 RT coordinates 0.101 and 0.391; similarities in coding strategy between
 RT insect and vertebrate iridoviruses";
 RL Virus Genes 15:235-245(1997).
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Delius H., Darai G., Fluegel R.M.;
 RT "DNA analysis of insect iridescent virus 6: evidence for circular
 RT permutation and terminal redundancy";
 RL J. Virol. 49:609-614(1984).
 RN [8]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=86174607; PubMed=3959991;
 RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
 RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
 RT in mice.";
 RL Med. Microbiol. Immunol. 175:43-53(1986).
 RN [9]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87321126; PubMed=2820141;
 RA Schnitzler P., Soltan J.B., Fischer M., Reisner H., Scholz J.,
 RA Delius H., Darai G.;
 RT "Molecular cloning and physical mapping of the genome of insect
 RT iridescent virus type 6: further evidence for circular permutation of
 RT the viral genome";
 RL Virology 160:66-74(1987).
 RN [10]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89073752; PubMed=3201750;
 RA Fischer M., Schnitzler P., Delius H., Darai G.;
 RT "Identification and characterization of the repetitive DNA element in
 RT the genome of insect iridescent virus type 6";
 RL Virology 167:485-496(1988).
 RN [11]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92196996; PubMed=1549908;
 RA Handermann M., Schnitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
 RA Darai G.;
 RT "Identification and mapping of origins of DNA replication within the
 RT DNA sequences of the genome of insect iridescent virus type 6";
 RL Virus Genes 6:19-32(1992).


```

RN  [12]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93260401; PubMed=8492091;
RA  Stohwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;
RT  "Identification of the gene encoding the major capsid protein of
RT  insect iridescent virus type 6 by polymerase chain reaction.";
RL  J. Gen. Virol. 74:873-879(1993).
RN  [13]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=94167241; PubMed=8121799;
RA  Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
RA  Delius H., Darai G.;
RT  "Identification of genes encoding zinc finger proteins, non-histone
RT  chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
RT  in the genome of Chilo iridescent virus.";
RL  Nucleic Acids Res. 22:158-166(1994).
RN  [14]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99383793; PubMed=10456793;
RA  Muller K., Tidona C.A., Darai G.;
RT  "Identification of a gene cluster within the genome of Chilo
RT  iridescent virus encoding enzymes involved in viral DNA replication
RT  and processing.";
RL  Virus Genes 18:243-264(1999).
RN  [15]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21342589; PubMed=11448171;
RA  Jakob N.J., Muller K., Bahr U., Darai G.;
RT  "Analysis of the first complete DNA sequence of an invertebrate
RT  iridovirus: coding strategy of the genome of Chilo iridescent virus.";
RL  Virology 286:182-196(2001).
DR  EMBL; AF303741; AAR82281.1; -.
SQ  SEQUENCE 40 AA; 4759 MW; 7EC07A3648016D9A CRC64;

Query Match 26.8%; Score 26; DB 2; Length 40;
Best Local Similarity 55.6%; Pred. No. 9.8e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 HLNSKIATFK 12
Db 29 NFNAKIALK 37

RESULT 153
AAR8804
ID AAR8804 PRELIMINARY; PRT; 40 AA.
AC AAR8804;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative chimeric protein 173-56.
OS Trypanosoma cruzi.
OC Eukaryota; Euzoenzoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC Schizotrypanum.
OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RA Teixeira A.R.L., Nitz N., Gomes C.;
RT "Germ line transmission of kDNA from Trypanosoma cruzi to mammals and
RT birds.";
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY490894; AAR8804.1; -.
SQ SEQUENCE 40 AA; 4736 MW; A4605A686F4B9C43 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 40;
Best Local Similarity 35.3%; Pred. No. 9.8e+03;
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 PNHLNSKIATFKVQSP 18
Db 18 PQFLNLRILKLYPPPLP 34

```

```

RESULT 154
Q7UQP8
ID Q7UQP8 PRELIMINARY; PRT; 41 AA.
AC Q7UQP8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=RB6178;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetaceae;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294143; CAD74652.1; -.
SQ SEQUENCE 41 AA; 4414 MW; 0EBBCAEFE28E305D CRC64;

Query Match 26.8%; Score 26; DB 2; Length 41;
Best Local Similarity 40.0%; Pred. No. 1e+04;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 PNHLNSKIATFK 11
Db 12 PNSFGGEVAF 21

RESULT 155
Q8F8H5
ID Q8F8H5 PRELIMINARY; PRT; 41 AA.
AC Q8F8H5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=UA0582;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AE011245; AAN47780.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 41 AA; 5144 MW; 02E1984E414596A5 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 41;
Best Local Similarity 36.4%; Pred. No. 1e+04;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 PNHLNSKIATFK 12
Db 10 PQYLNKXFLYQ 20

```

RT burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AE000787; AAC66101.1; -.
DR PIR; H70246; H70246.
DR TIGR; BBJ15; -.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 42 AA; 4866 MW; B02DD8EC5DB265CC CRC64;

Query Match 26.8%; Score 26; DB 2; Length 42;
Best Local Similarity 62.5%; Pred. No. 1e+04; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 2;

QY 9 IAFKIVSQ 16
| | | | |
DB 8 IIFKFISQ 15

RESULT 158
Q73I36 PRELIMINARY; PRT; 42 AA.
ID Q73I36
AC Q73I36;
DT 05-JUN-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=WD0347;
OS Wolbachia pipientis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15024419;
RA Wu M., Sun L.V., Vamathevan J.J., Riegler M., DeBoy R.T.,
RA Brownlie J.C., McGraw E.A., Martin W., Esser C., Ahmadinejad N.,
RA Wiegand C., Madupu R., Beanan M.O., Brinkac L.M., Daugherty S.C.,
RA Durkin A.S., Kolonay J.F., Nelson W.C., Mohamoud Y., Lee P.,
RA Berry K.J., Young M.B., Utterback T.R., Weidman J.F., Nierman W.C.,
RA Paulsen I.T., Nelson K.E., Tettelin H., O'Neill S.I., Eisen J.A.;
RT "Phylogenomics of the reproductive parasite Wolbachia pipientis wMel:
RT a streamlined genome overrun by mobile genetic elements.";
RL PLOS Biol. 2:327-341(2004).
DR EMBL; AE017257; AAS14076.1; -.
DR TIGR; WD0347; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 42 AA; 4976 MW; 6C42B5AE06425A06 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 42;
Best Local Similarity 42.9%; Pred. No. 1e+04; Indels 0; Gaps 0;
Matches 3; Conservative 3; Mismatches 1;

QY 1 EPNHINS 7
| | | | |
DB 24 DPNHIXN 30

RESULT 159
AAS14076 PRELIMINARY; PRT; 42 AA.
ID AAS14076
AC AAS14076;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 13-APR-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN WD0347
OS Wolbachia pipientis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15024419;
RA Wu M., Sun L.V., Vamathevan J.J., Riegler M., DeBoy R.T.,

Q33617 PRELIMINARY; PRT; 42 AA.
ID Q33617
AC Q33617;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Competence stimulating peptide precursor.
GN Name=cmC;
OS Streptococcus cristatus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=45634;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 12479;
RX MEDLINE=98012953; PubMed=9352904;
RA Havarstein L.S., Hakenbeck R., Gaustad P.;
RT "Natural competence in the genus Streptococcus: Evidence that
RT streptococci can change phenotype by interspecies recombinational
RT exchanges.";
RL J. Bacteriol. 179:6589-6594(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 12479;
RA Havarstein L.S., Hakenbeck R., Gaustad P.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ000876; CAA04366.1; -.
DR GO; GO:0005186; F:pheromone activity; IEA.
DR InterPro; IPR004288; ComC.
DR Pfam; PF03047; ComC; 1.
KW Signal.
FT SIGNAL 1 28 Potential.
FT CHAIN 29 42 competence stimulating peptide.
SQ SEQUENCE 42 AA; 4979 MW; 8EF35127D86FAABD CRC64;

Query Match 26.8%; Score 26; DB 2; Length 42;
Best Local Similarity 38.5%; Pred. No. 1e+04; Indels 0; Gaps 0;
Matches 5; Conservative 4; Mismatches 4;

QY 6 NSKIAFKIVSQEP 18
| | | | |
DB 3 NTKNFXTIAQFP 15

RESULT 157
O50770 PRELIMINARY; PRT; 42 AA.
ID O50770
AC O50770;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein BBJ15.
GN OrderedLocusNames=BBJ15;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp38.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403695; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.P., Utterback T.R., Watthey L., McDonald L.A.,
RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia

```

RA Brownlie J.C., McGraw E.A., Martin W., Esser C., Ahmadinejad N.,
RA Wiegand C., Madupu R., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Durkin A.S., Kolonay J.F., Nelson W.C., Mohammad Y., Lee P.,
RA Berry K.J., Young M.B., Utterback T.R., Weidman J.F., Nierman W.C.,
RA Paulsen I.F., Nelson K.E., Tettelin H., O'Neill S.L., Eisen J.A.,
RA "Phylogenomics of the reproductive parasite Wolbachia pipientis wMel:
RA a streamlined genome overrun by mobile genetic elements.";
RL PLOS Biol. 2:327-341(2004).
DR ENBL; AE017257; AAS14076.1; -.
DR TIGR; WD0347; -.
KW Hypothetical protein.
SQ SEQUENCE 42 AA; 4976 MW; 6C42B5AE06425A06 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 42;
Best Local Similarity 42.9%; Pred. No. 1e+04;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLS 7
Db :|||:
24 DPNHKN 30

RESULT 160
O43456 PRELIMINARY; PRT; 43 AA.
AC O43456;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97239231; PubMed=9084893;
RA Yee F., Yolken R.H.;
RT "Identification of differentially expressed RNA transcripts in
RT neuropsychiatric disorders.";
RL Biol. Psychiatry 41:759-761(1997).
DR ENBL; AF017336; AAB92502.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 43 AA; 5119 MW; 720F6E2726EE005B CRC64;

Query Match 26.8%; Score 26; DB 2; Length 43;
Best Local Similarity 57.1%; Pred. No. 1e+04;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 HNSKIA 10
Db :|||:
7 HNRKLA 13

RESULT 161
Q7MECO PRELIMINARY; PRT; 43 AA.
AC Q7MECO;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein VWA0750.
GN Name=VWA0750;
OS Vibrio vulnificus (strain Y016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656945;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Shang A.B.-T., Li J.-C., Su T.-L.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Chen A.B.-T., Li J.-C., Su T.-L.,

```

```

RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RA "Comparative genome analysis of Vibrio vulnificus, a marine
RA pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR ENBL; AP005347; BAC96776.1; -.
KW Hypothetical protein.
SQ SEQUENCE 43 AA; 5002 MW; 9F0BA3252600C257 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 43;
Best Local Similarity 44.4%; Pred. No. 1e+04;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 NHLNSKIAF 11
Db :|||:
30 HHLNAAVYF 38

RESULT 162
Q7UH93 PRELIMINARY; PRT; 43 AA.
AC Q7UH93;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RB4772;
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR ENBL; BX294141; CAD78083.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 43 AA; 4939 MW; F39C911A65E638F4 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 43;
Best Local Similarity 71.4%; Pred. No. 1e+04;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 SKIAFKI 13
Db :|||:
30 SKVAFNI 36

RESULT 163
Q9KLP2 PRELIMINARY; PRT; 43 AA.
AC Q9KLP2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VCA0701.
GN OrderedLocusNames=VCA0701;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,

```

DE Hypothetical protein.
GN OrderedLocusNames=LB315;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
interrogans revealed by whole-genome sequencing."
RL Nature 422:888-893(2003).
DR EMBL; AE011618; AAN51874.1; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 44 AA; 5110 MW; FE15B5EF2893C73B CRC64;

Query Match 26.8%; Score 26; DB 2; Length 44;
Best Local Similarity 71.4%; Pred. No. 1.1e+04;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 9 IAFKIVS 15
| | | | |
Db 25 IAFQVVS 31

RESULT 166
Q8EZK9 PRELIMINARY; PRT; 44 AA.
ID Q8EZK9
AC Q8EZK9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LA3844;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
interrogans revealed by whole-genome sequencing."
RL Nature 422:888-893(2003).
DR EMBL; AE011539; AAN51042.1; -;
DR InterPro; IPR011461; DUF1567.
DR Pfam; PF07604; DUF1567; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 44 AA; 5265 MW; CF36FD3549F6F165 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 44;
Best Local Similarity 57.1%; Pred. No. 1.1e+04;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNHLNSK 8
| | | | |
Db 38 PAHVNNK 44

RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D., R.,
Nierman W.C., White O., Salzberg S.B., Smith H.O., Colwell R.R.,
Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae";
RL Nature 406:477-483(2000).
DR EMBL; AE004399; AAF96600.1; -;
DR PIR; E82428; E82428.
DR TIGR; VCA0701; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 43 AA; 5088 MW; 3F1DC3748E1BEFCF CRC64;

Query Match 26.8%; Score 26; DB 2; Length 43;
Best Local Similarity 35.7%; Pred. No. 1.1e+04;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQEP 18
| | | | |
Db 24 LYSKVLRKILSEP 37

RESULT 164
Q9G665 PRELIMINARY; PRT; 44 AA.
ID Q9G665
AC Q9G665;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NADH dehydrogenase subunit 2.
GN Name=ND2;
OS Diporiphora bilineata (Two-lined dragon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Amphisbolurinae;
OC Diporiphora.
OX NCBI_TaxID=118204;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22114082; PubMed=12118408;
RA Macey J.R., Schulte J.A. II, Larson A., Ananjeva N.B., Wang Y.,
Pethiyagoda R., Rastegar-Pouyani N., Papenfuss T.J.;
RT "Evaluating trans-tethys migration: an example using acrodont lizard
phylogenetics";
RL Syst. Biol. 49:233-256(2000).
DR EMBL; AF128473; AAG00655.2; -;
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
SQ SEQUENCE 44 AA; 4390 MW; AB6E44A167C25B6D CRC64;

Query Match 26.8%; Score 26; DB 2; Length 44;
Best Local Similarity 44.4%; Pred. No. 1.1e+04;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PNHLNSKIA 10
| | | | |
Db 36 PRHTNSHLS 44

RESULT 165
Q8EX99 PRELIMINARY; PRT; 44 AA.
ID Q8EX99
AC Q8EX99;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

```

RESULT 167
Db Q8VJ61 PRELIMINARY; PRT; 44 AA.
ID Q8VJ61
AC Q8VJ61
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein MT3273.
GN OrderedLocusNames=MT3273;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AB007140; AAK47613.1; -.
DK TIGR; MT3273; -.
KW Hypothetical protein.
SQ SEQUENCE 44 AA; 4759 MW; E4410301A7E125ED CRC64;

Query Match 26.8%; Score 26; DB 2; Length 44;
Best Local Similarity 35.7%; Pred. No. 1.1e+04;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 HLNSKIAFKIVSQE 17
Db 5 HLPSKLGHPKVLRK 18

RESULT 168
ID Q85413 PRELIMINARY; PRT; 44 AA.
AC Q85413
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nucleoprotein N (Fragment).
OS Rabies virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11292;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=84216456; PubMed=6328006;
RA Kurilla M.G., Cabradilla C.D., Holloway B.P., Keene J.D.;
RT "Nucleotide sequence and host La protein interactions of rabies virus
RT leader RNA.";
RL J. Virol. 50:773-778(1984).
DR EMBL; M12771; AAA47221.1; -.
DR GO; GO:0019013; C:Viral nucleocapsid; IEA.
DR InterPro; IPR000448; Rhabd_nucleocap.
DR Pfam; PF00945; Rhabdo_ncap; 1.
DR ProDom; PD002087; Rhabd_nucleocap; 1.
KW Nucleocapsid.
FT NON_TER 44
SQ SEQUENCE 44 AA; 5084 MW; 4539CCFC4C24D6BC CRC64;

Query Match 26.8%; Score 26; DB 2; Length 44;
Best Local Similarity 55.6%; Pred. No. 1.1e+04;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 KIAFKIVSQ 16

```

```

Db 5 XIVEKVNQ 13

RESULT 169
Q7LZ30 PRELIMINARY; PRT; 44 AA.
ID Q7LZ30
AC Q7LZ30
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein 1 (Fragment).
OS Coelognathus radiatus (Radiated ratsnake) (Elaphe radiata).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Colubridae; Colubrinae; Coelognathus.
OX NCBI_TaxID=201391;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92247938; PubMed=6954524;
RX Eppelen J.F., McCarrey J.R., Sutou S., Ohno S.;
RA "Base sequence of a cloned snake W-chromosome DNA fragment and
RT identification of a male-specific putative mRNA in the mouse.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3798-3802(1982).
DR PIR; A19434; A19434.
FT NON_TER 1
FT NON_TER 44
SQ SEQUENCE 44 AA; 4928 MW; B8E37434B8DB2B80 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 44;
Best Local Similarity 57.1%; Pred. No. 1.1e+04;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNFHNSK 8
Db 11 PSHLKS 17

RESULT 170
Q7S7W6 PRELIMINARY; PRT; 45 AA.
ID Q7S7W6
AC Q7S7W6
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Predicted protein.
GN Name=NCU01316.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
RA Seditrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Roth G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysellis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDbJ whole genome shotgun (WGS) entry which is
CC preliminary data.

```

```
DR EMBL; AABX01000269; EAA32166.1; -.
SQ SEQUENCE 45 AA; 5090 MW; DEFAB082E2CD192D CRC64;

Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 50.0%; Pred. No. 1.1e+04;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 SKIAFKIVSQ 16
Db 14 AKYKYKIVSE 23

RESULT 171
Q9P1G5 PRELIMINARY; PRT; 45 AA.
AC Q9P1G5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PRO1828.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA Liu M., He F.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF116669; AAF71089.1; -.
SQ SEQUENCE 45 AA; 4643 MW; FA43813FD107F6F5 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 38.5%; Pred. No. 1.1e+04;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 7 SKIAFKIVSQEPA 19
Db 11 SEVATKVFQASPA 23

RESULT 172
QYVF4 PRELIMINARY; PRT; 45 AA.
AC QYVF4
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=tb927.2.2780;
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22789168; PubMed=12907728;
RA El-Sayed N.M.A., Ghedin E., Song J., MacLeod A., Bringaud F.,
RA Larkin C., Wanless D., Peterson J., Hou L., Taylor S., Tweedie A.,
RA Biteau N., Khalak H.G., Lin X., Mason T., Hannick L., Caler E.,
RA Blandin G., Bartholomeu D., Simpson A.J., Kaul S., Zhao H., Pai G.,
RA Van Aken S., Utterback T., Haas B., Koo H.L., Umayam L., Suh B.,
RA Gerrard C., Leech V., Qi R., Zhou S., Schwartz D., Feldblyum T.,
RA Salzberg S., Tait A., Turner M.R., Ullu E., White O., Melville S.,
RA Adams M.D., Fraser C.M., Donelson J.E.;
RT "The sequence and analysis of Trypanosoma brucei chromosome II.";
RL Nucleic Acids Res. 31:4856-4863(2003).
DR EMBL; AF017168; RAQ15735.1; -.
KW Hypothetical protein.
SQ SEQUENCE 45 AA; 5511 MW; 432CD1DC4EB80F63 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 40.0%; Pred. No. 1.1e+04;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 HLNSKIAFKI 13
Db 11 HMNFSIVYEI 20

RESULT 173
Q8IG46 PRELIMINARY; PRT; 45 AA.
AC Q8IG46
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C25H3.6.
GN Names=C25H3.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Johnson D., Wilson R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; U29535; AAN63450.3; -.
DR WormPep; C25H3.6C; CE32141.
KW Hypothetical protein.
SQ SEQUENCE 45 AA; 4701 MW; E2814B9A9640A2AB CRC64;

Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 55.6%; Pred. No. 1.1e+04;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNHLSKITA 10
Db 13 PNHLSVSWA 21

RESULT 174
Q7M214 PRELIMINARY; PRT; 45 AA.
AC Q7M214
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major acute phase protein (Fragments).
OS Solanum crispum (Chilean potato-tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
```

```

OC lamiids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4110;
RN [1]
RP SEQUENCE.
RA Gonzalez-Ramon N., Alava M.A., Sarsa J.A., Pineiro M., Escartin A.,
RA Garcia-Gil A., Lampreave F., Pineiro A.;
RT "The major acute phase serum protein in pigs is homologous to human
RT plasma kallikrein sensitive PK-120.";
RL FEBS Lett. 371:227-230(1995).
DR PIR; S66648; S66648.
FT NON_TER 1
FT NON_TER 45
FT NON_TER 45
SQ SEQUENCE 45 AA; 5066 MW; A9846777D5674618 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 28.6%; Pred. No. 1.1e+04;
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 5 LNSKIAPKIVSQEP 18
   : : : : :
Db 16 VSRFAKVVTSSTP 29

RESULT 175
Q9SAM1 PRELIMINARY; PRT; 45 AA.
AC Q9SAM1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE R2R3-MYB transcription factor (Fragment).
GN Name=AtMYB73;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole plants;
RA Romero I., Fuentes A., Benito M.J., Malpica J., Leyva A., Paz-Ares J.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -|- SIMILARITY: Contains 1 Myb-like domain.
DR EMBL; Z95796; CAB09228.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR009057; Homeodomain-like.
DR Pfam; PF00249; Myb DNA binding.
DR PROSITE; PS50090; MYB_3; 1.
KW Nuclear protein.
FT NON_TER 1
FT NON_TER 45
FT NON_TER 45
SQ SEQUENCE 45 AA; 5337 MW; 68249D3ECBE4051D CRC64;

Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 33.3%; Pred. No. 1.1e+04;
Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 3 NHLSKIAPKIVSQE 17
   : : : : :
Db 5 NQLSPVEHRAFSQE 19

RESULT 176
Q7ULY8 PRELIMINARY; PRT; 45 AA.
AC Q7ULY8
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.

```

```

CN OrderedLocustNames=RB9191;
OS Rhodopirellula Baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294149; CAD76129.1; -.
DR InterPro; IPR011479; DUF1586.
DR Pfam; PF07625; DUF1586; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 45 AA; 4582 MW; 8C0D405930F78473 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 50.0%; Pred. No. 1.1e+04;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SKIAFKIVSQEP 18
   : : : : :
Db 2 SRTALAAVSQTP 13

RESULT 177
Q83NY5 PRELIMINARY; PRT; 45 AA.
AC Q83NY5
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative integral membrane protein.
GN OrderedLocustNames=TW024;
OS Tropheryma whippelii (strain TW08/27) (Whipple's bacillus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
OX NCBI_TaxID=218496;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22495039; PubMed=12606174;
RA Bentley S.D., Maiwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,
RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,
RA Barrell B.G., Parkhill J., Rellman D.A.;
RT "Sequencing and analysis of the genome of the Whipple's disease
RT bacterium Tropheryma whippelii.";
RL Lancet 361:637-644(2003).
DR EMBL; BX251410; CAD66716.1; -.
KW Complete proteome.
SQ SEQUENCE 45 AA; 5155 MW; 1356DB39D8D00C70 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 50.0%; Pred. No. 1.1e+04;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 9 IAFKIVSQEP 18
   : : : : :
Db 24 IAFTVAVVSP 33

RESULT 178
AAAG3450 PRELIMINARY; PRT; 45 AA.
ID AAAG3450
AC AAAG3450;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)

```

DE Hypothetical protein C25H3.6.
GN C25H3.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Johnson D., Wilson R.;
RA "the sequence of C. elegans cosmid C25H3.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Waterston R.;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Wilson R.;
RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Wilson R.;
RA Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; U29535; AAN63450.3; -.
KW Hypothetical protein.
SQ SEQUENCE 45 AA; 4701 MW; E2814B9A9640A2AB CRC64;
Query Match 26.8%; Score 26; DB 2; Length 45;
Best Local Similarity 55.6%; Pred. No. 1.1e+04;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 PNHLSKJIA 10
Db 13 PNHLSVMA 21
[1]
RESULT 179
AOFA SHEEP STANDARD; PRT; 46 AA.
AC O18851;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Amine oxidase [flavin-containing] A (EC 1.4.3.4) (Monoamine oxidase)
DE (MAO-A) (Fragment).
GN Name=MAOA;
OS Ovis aries (Sheep).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=98251499; PubMed=9589594;
RX Cambridge L.M., Lumsden J.M., Sadhigi M., Galloway S.M.;
RA "A Scal polymorphism at the ovine monoamine oxidase A locus (MAOA).";
RL Anim. Genet. 28:457-457(1997).
CC -!- FUNCTION: Catalyzes the oxidative deamination of biogenic and
CC xenobiotic amines and has important functions in the metabolism of
CC neuroactive and vasoactive amines in the central nervous system
CC and peripheral tissues. MAO-A preferentially oxidizes biogenic

CC amines such as 5-hydroxytryptamine (5-HT), norepinephrine and
CC epinephrine (By similarity).
CC -!- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC H(2)O(2). FAD.
CC -!- COFACTOR: FAD.
CC -!- SUBUNIT: Monomer, homo- or heterodimer (containing two subunits of
CC similar size). Each subunit contains a covalently bound flavin.
CC Enzymatically active as monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Mitochondrial outer membrane.
CC -!- SIMILARITY: Belongs to the flavin monoamine oxidase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF023613; AAC16911.1; -.
CC HSP; P27338; IGOS.
CC Catecholamine metabolism; FAD; Flavoprotein; Mitochondrion;
CC Neurotransmitter degradation; Oxidoreductase; Transmembrane.
CC FT NON TER 1
CC 46
CC SQ SEQUENCE 46 AA; 5349 MW; FD06EBDE44F88BA CRC64;
Query Match 26.8%; Score 26; DB 1; Length 46;
Best Local Similarity 54.5%; Pred. No. 1.1e+04;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 PNHLSKIAFK 12
Db 7 PPTLTSKIHR 17
[1]
RESULT 180
Q7PJP3 PRELIMINARY; PRT; 46 AA.
AC Q7PJP3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000024134.
GN Name=ENSANG0000020361;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008978; EAA43681.1; -.
SQ SEQUENCE 46 AA; 5369 MW; EDC1921A05DFD672 CRC64;
Query Match 26.8%; Score 26; DB 2; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+04;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 9 IAFKIVSQ 16
Db 26 IAFKUTQQ 33
[1]
RESULT 181
Q9LA74 PRELIMINARY; PRT; 46 AA.
ID Q9LA74
AC Q9LA74;
[1]

RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serrero P.,
 RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
 RA Takemura K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
 RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
 RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
 RA Weltzenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
 RA Yoshikawa H., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
 RA "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis";
 RL Nature 390:249-256(1997).
 RN [2]
 RP DEVELOPMENTAL STAGE, AND REGULATION OF EXPRESSION.
 RX MEDLINE=99267404; PubMed=10333516; DOI=10.1016/S0378-1119(99)00124-9;
 RA Cabrera-Hernandez A., Sanchez-Salas J.-L., Paldhungat M., Setlow P.,
 RT "Regulation of four genes encoding small, acid-soluble spore proteins
 RT in Bacillus subtilis";
 RL Gene 232:1-10(1999).
 CC -!- SUBCELLULAR LOCATION: Spore core (Probable).
 CC -!- DEVELOPMENTAL STAGE: Expressed only in the forespore compartment
 CC of sporulating cells.
 CC -!- INDUCTION: Expression is sigma F and sigma G-dependent.
 CC -!- SIMILARITY: Belongs to the spn family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z99113; CAE01455.1; -;
 DR Subtilist; BG14179; sspN.
 DR HAMAP; MF_01505; -; 1.
 KW Complete proteome; Sporulation.
 SQ SEQUENCE 48 AA; 5353 MW; 283A62D662070859 CRC64;

 Query Match 26.8%; Score 26; DB 1; Length 48;
 Best Local Similarity 57.1%; Pred. No. 1.2e+04;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

 Qy 2 PHLNSK 8
 Db 14 PSHLGTK 20
 |::|:|

 RESULT 183
 Q15455 PRELIMINARY; PRT; 48 AA.
 ID Q15455
 AC Q15455;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Protein-serine/threonine kinase (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RL [1]
 RP SEQUENCE FROM N.A.
 RN MEDLINE=94100173; PubMed=8274451;
 RX Schultz S.J., Nigg E.A.;
 RA "Identification of 21 novel human protein kinases, including 3 members
 RT of a family related to the cell cycle regulator nima of Aspergillus
 RT nidulans";
 RT Cell Growth Differ. 4:821-830(1993).
 RL Cell Growth Differ. 4:821-830(1993).
 RL EMBL; Z25433; CAA80920.1; -;
 DR PIR; I38223; I38223.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0006469; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.

DR InterPro: IPR000719; Prot_kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Kinase.
FT NON_TER 1 1
FT NON_TER 48 48
SQ SEQUENCE 48 AA; 5319 MW; 3CF812F406559080 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 48;
Best Local Similarity 44.4%; Pred. No. 1.2e+04;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNHLNSKIA 10
Db 25 PNVISPEIA 33
| | | | |
| | | | |

RESULT 184
Q85WS4 PRELIMINARY; PRT; 48 AA.
AC Q85WS4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF48a.
OS Pinus koraiensis (Korean pine).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=88728;
RN [1]
RP SEQUENCE FROM N.A.
RA Noh E.W., Lee J.S., Choi Y.I., Han M.S., Yi Y.S., Han S.U.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY228468; AA074145.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
SQ SEQUENCE 48 AA; 5459 MW; B10D5EA3528FC44 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 48;
Best Local Similarity 38.5%; Pred. No. 1.2e+04;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 6 NSKIAFKIVSQRP 18
Db 15 HSNVPFRHVCTEP 27
: | : | : | : |
: | : | : | : |

RESULT 185
Q9X472 PRELIMINARY; PRT; 48 AA.
AC Q9X472;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE EntF.
GN Name=entF;
OS Enterococcus faecium (Streptococcus faecium).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1352;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DPC1146;
RX MEDLINE=99203104; PubMed=10103244;
RA O'Keefe T., Hill C., Ross R.P.;
RT "Characterization and heterologous expression of the genes encoding enterocin A production, immunity, and regulation in Enterococcus faecium DPC1146;"
RT faecium DPC1146;"
RL Appl. Environ. Microbiol. 65:1506-1515(1999).
DR EMBL; AF09088; AA029134.1; -.
SQ SEQUENCE 48 AA; 5297 MW; 971753DE7BB56A4D CRC64;

Query Match 26.8%; Score 26; DB 2; Length 48;
Best Local Similarity 62.5%; Pred. No. 1.2e+04;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNELNSK 8
Db 3 EKURLNAK 10
| | | | |
| | | | |

RESULT 186
Q8XFZ3 PRELIMINARY; PRT; 48 AA.
AC Q8XFZ3; Q7ALT3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein STY4474.
GN OrderedLocusNames=STY4474, t4182;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and Cr18;"
RL J. Bacteriol. 185:2330-2337(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Cr18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1039/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi Cr18;"
RL Nature 413:848-852(2001).
DR EMBL; AF016848; AA071646.1; -.
DR EMBL; AL627282; CAD09260.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 48 AA; 5705 MW; 703E0FC783E036C4 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 48;
Best Local Similarity 62.5%; Pred. No. 1.2e+04;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NSKIAFKI 13
Db 24 NSKVVFYI 31
| | | | |
| | | | |

RESULT 187
Q7CPA7 PRELIMINARY; PRT; 48 AA.
AC Q7CPA7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative cytoplasmic protein.
GN OrderedLocusNames=STM4276;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;

```

RN SEQUENCE FROM N.A.
RP STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008900; AAL23100.1; -.
KW Complete proteome.
SQ SEQUENCE 48 AA; 5705 MW; 703E0FC783E036C4 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 48;
Best Local Similarity 62.5%; Pred. No. 1.2e+04;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NSKIAFKI 13
Db 24 NSKVVEYI 31

RESULT 188
Q8AFGI ID Q8AFGI PRELIMINARY; PRT; 48 AA.
AC Q8AFGI
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Envelope polyprotein (Fragment).
OS Mouse mammary tumor virus.
OC Viruses; Retroid viruses; Retroviridae; Betaretrovirus;
OC Mammalian type B retroviruses.
OX NCBI_TaxID=11757;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=10016;
RA Polesz B.J., Perzova R.N., Abbott L.Z.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY152723; AAN7719.1; -.
DR GO; GO:0019031; C:Viral envelope; IEA.
KW Envelope protein; Polyprotein.
FT NON TER 1
FT NON TER 1
FT NON TER 48
SQ SEQUENCE 48 AA; 5402 MW; 67028A527E8DF382 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 48;
Best Local Similarity 45.5%; Pred. No. 1.2e+04;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PNHLSKIAFK 12
Db 6 PKYPHCQIAFK 16

RESULT 189
Q9UHS1 ID Q9UHS1 PRELIMINARY; PRT; 49 AA.
AC Q9UHS1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PRO2086.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;

```

```

RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Bi J., Zhang Y., Liu M.,
RA He F.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF118093; AAF22037.1; -.
SQ SEQUENCE 49 AA; 5519 MW; 84E0A1CE6687C2B2 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 49;
Best Local Similarity 54.5%; Pred. No. 1.2e+04;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVS 15
Db 16 LNEXYIFKKVA 26

RESULT 190
Q8HYK6 ID Q8HYK6 PRELIMINARY; PRT; 49 AA.
AC Q8HYK6
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Heart fatty acid-binding protein (Fragment).
GN Name=HFABP;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX Calvo J.H., Serrano M., Jurado J.J.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF473547; AAN86782.1; -.
DR HSP; P10790; IBWY.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011038; Calycin.
DR InterPro; IPR000566; Lipoclin_cytFABP.
DR Pram; PF00061; Lipocalin; 1.
FT NON TER 1
FT NON TER 49
FT NON TER 49
SQ SEQUENCE 49 AA; 5437 MW; 0B74BD8554ED86C3 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 1.2e+04;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 NSKIAFKI 13
Db 33 NTEISPKL 40

RESULT 191
Q8CLI5 ID Q8CLI5 PRELIMINARY; PRT; 49 AA.
AC Q8CLI5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical.
GN OrderedLocusNames=y1269;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,

```

```
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL: AE013730; AM84844.1; -.
KW Hypothetical protein.
SQ SEQUENCE 49 AA; 5454 MW; 7490D4C0D71D0E4D CRC64;

Query Match 26.8%; Score 26; DB 2; Length 49;
Best Local Similarity 26.3%; Pred. No. 1.2e+04;
Matches 5; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 17 EPRLPANVNFVAGKPPA 35

RESULT 192
P88744 PRELIMINARY; PRT; 49 AA.
AC P88744;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nef protein (Fragment).
GN Name=nef;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1];
RP SEQUENCE FROM N.A.
RA Quinones-Mateu M.E., Domingo E.;
RL Submitted (DRC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen (by
CC similarity).
DR EMBL: U81469; AB39173.1; -.
DR GO: GO:0005525; F:GTP binding; IEA.
DR InterPro: IPR000008; C2.
DR InterPro: IPR001558; HIV Nef.
DR Pfam: PF00469; F-protein; 1.
DR PRINTS: PR00360; C2DOMAIN.
DR ProDom: PD000031; HIV Nef; 1.
DR ProDom: PD000031; HIV Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER 1
SQ SEQUENCE 49 AA; 5768 MW; 2EC4DF5E0AB09976 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 49;
Best Local Similarity 41.7%; Pred. No. 1.2e+04;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NSKIAPKIVSQE 17
Db 33 DSRLAFHHVARE 44

RESULT 193
Q7RMX5 PRELIMINARY; PRT; 50 AA.
AC Q7RMX5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY02053;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteu M.,
```

```
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shalom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AABL01000563; EAA21471.1; -.
KW Hypothetical protein.
SQ SEQUENCE 50 AA; 6018 MW; 9D2374C6C8A84E61 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 50;
Best Local Similarity 46.2%; Pred. No. 1.2e+04;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKI 13
Db 24 DPNTENKKYGGKI 36

RESULT 194
Q55072 PRELIMINARY; PRT; 50 AA.
ID Q55072;
AC Q55072;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein (Fragment).
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE=97449302; PubMed=9305771;
RA Cassier-Chauvat C., Poncelet M., Chauvat F.;
RT "Three insertion sequences from the cyanobacterium Synechocystis
RT PCC6803 support the occurrence of horizontal DNA transfer among
RT bacteria.";
RL Gene 195:257-266(1997).
DR EMBL: U38915; AAB72128.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 50 AA; 5711 MW; 637921F14554C893 CRC64;

Query Match 26.8%; Score 26; DB 2; Length 50;
Best Local Similarity 31.6%; Pred. No. 1.2e+04;
Matches 6; Conservative 6; Mismatches 5; Indels 2; Gaps 1;

Qy 3 NHHLSKIAPKI--VSOEPA 19
Db 28 NHENNAVVDIEQKLAQNF 46

RESULT 195
Q9JVI7 PRELIMINARY; PRT; 50 AA.
ID Q9JVI7;
AC Q9JVI7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein NMA0824.
GN OrderedLocusNames=NMA0824;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
```

```

OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Skellon J.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506 (2000).
DR EMBL; AL162754; CAB84106.1; --
DR PIR; E81927; E81927.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 50 AA; 5434 MW; 519D26F56A81BA4A CRC64;

Query Match 26.8%; Score 26; DB 2; Length 50;
Best Local Similarity 66.7%; Pred. No. 1.2e+04;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 10 AFKIVSQEP 18
Db ||| ||| |||
15 AFKPVLRQP 23

RESULT 196
ID O61191 PRELIMINARY; PRT; 37 AA.
AC O61191;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histone H3 (Fragment).
GN Name=H3-1;
OS Blepharisma undulans.
OC Eukaryota; Alveolata; Ciliophora; Heterotricha; Heterotrichida;
OC Blepharismidae; Blepharisma.
OX NCBI_TaxID=74723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98154410; PubMed=9493359;
RA Bernhard D., Schlegel M.;
RT "Evolution of histone H4 and H3 genes in different ciliate lineages."
RL J. Mol. Evol. 46:344-354 (1998).
DR EMBL; AJ004689; CAA06039.1; --
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0007001; P:chromosome organization and biogenesis (sen. .; IEA.
DR InterPro; IPR009072; Histone-fold.
DR PRINTS; PR00622; HISTONEH3.
FT NON TER 37
SQ SEQUENCE 37 AA; 3884 MW; B7F723ECF3F9C6F2 CRC64;

Query Match 26.3%; Score 25.5; DB 2; Length 37;
Best Local Similarity 56.2%; Pred. No. 1.1e+04;
Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Oy 4 HLNSKIAFKIVSQEPA 19
Db ||| ||| ||| |||
21 HLAATKAA-KKVSQTTA 35

RESULT 197
Q820D2
ID Q8E0D2 PRELIMINARY; PRT; 38 AA.
AC Q8E0D2;

01-MAR-2003 (TrEMBLrel. 23, Created)
01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Hypothetical protein SAG0802.
OrderedLocusNames=SAG0802;
Streptococcus agalactiae (serotype V).
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
NCBI_TaxID=216466;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
RA Tectelin H., Masignani V., Cieslewicz M.J., Eissen J.A., Peterson S.N.,
RA Wessels L.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Pedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae."
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AE014228; AAM99689.1; --
DR TIGR; SAG0802; --
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 38 AA; 4578 MW; AB784C66B92E3E18 CRC64;

Query Match 26.3%; Score 25.5; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 1.1e+04;
Matches 6; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Oy 5 LNSKI-AFKIVS 15
Db ||| : ||| :
4 LNSMVPFSKIIT 15

RESULT 198
YPI_BPPRD
ID YPI_BPPRD STANDARD; PRT; 42 AA.
AC P27386;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical 4.5 kDa protein in genes IX-III intergenic region.
GN Name=I;
OS Bacteriophage PRD1.
OC Viruses; dsDNA viruses, no RNA stage; Tectiviridae; Tectivirus.
OX NCBI_TaxID=10658;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91306449; PubMed=1853567; Pakula T.M., Ojala P.M.,
RA Bamford J.K.H., Haenninen A.-L., Bamford D.H.;
RA Kalkkinen N., Frilander M., Bamford D.H.;
RT "Genome organization of membrane-containing bacteriophage PRD1."
RL Virology 183:658-676 (1991).
CC -!- FUNCTION: Either one of the ORFs (I and J) or both of them form
CC the gene XX for protein P20 (responsible for DNA packaging).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M69077; AAA32464.1; --
DR PIR; A36776; WNEPPI.
KW Hypothetical protein.

```

SQ	SEQUENCE	42 AA; 4541 MW; B89A83DED4FFA8FA CRC64;			
Query Match	26.3%;	Score 25.5; DB 1; Length 42;			
Best Local Similarity	42.9%;	Pred. No. 1.3e+04;			
Matches	6; Conservative	4; Mismatches	1; Indels	3; Gaps	1;
Qy	9 IAFKIV---SOEPA 19	: : : :			
Db	29 LAKVINNRAQEPA 42	: : : :			
RESULT 199					
Q6EDX2	PRELIMINARY;	PRT; 42 AA.			
AC	Q6EDX2;				
DT	01-OCT-2004 (TREMBlrel. 28, Created)				
DT	01-OCT-2004 (TREMBlrel. 28, Last sequence update)				
DT	01-OCT-2004 (TREMBlrel. 28, Last annotation update)				
DE	ORFi.				
OS	Bacteriophage PR772.				
OC	Viruses; dsDNA viruses, no RNA stage; Tectiviridae; Tectivirus.				
OX	NCBI_TaxID=261665;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Lute S.C., Aranha H., Tremblay D., Liang D., Ackermann H.-W., Chu B.,				
RA	Molineau S., Brorson K.A.;				
RL	Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY441783; AAR99752.1; -.				
SQ	SEQUENCE 42 AA; 4541 MW; B89A83DED4FFA8FA CRC64;				
Query Match	26.3%;	Score 25.5; DB 2; Length 42;			
Best Local Similarity	42.9%;	Pred. No. 1.3e+04;			
Matches	6; Conservative	4; Mismatches	1; Indels	3; Gaps	1;
Qy	9 IAFKIV---SOEPA 19	: : : :			
Db	29 LAKVINNRAQEPA 42	: : : :			
RESULT 200					
Q8F8F3	PRELIMINARY;	PRT; 44 AA.			
AC	Q8F8F3;				
DT	01-MAR-2003 (TREMBlrel. 23, Created)				
DT	01-MAR-2003 (TREMBlrel. 23, Last sequence update)				
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)				
DE	Hypothetical protein.				
GN	OrderedlocusNames=LA0604;				
OS	Leptospira interrogans.				
OC	Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.				
OX	NCBI_TaxID=173;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;				
RX	MEDLINE=22598143; PubMed=12712204;				
RA	Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,				
RA	Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,				
RA	Jiang Y.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,				
RA	Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,				
RA	Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,				
RA	Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,				
RA	Xu J.-G., Zhao G.-P.;				
RT	"Unique physiological and pathogenic features of Leptospira				
RT	interrogans revealed by whole-genome sequencing.";				
RL	Nature 422:888-893 (2003).				
DR	EMBL; AF011247; AAN47803.1; -.				
KW	Complete proteome; Hypothetical protein.				
SQ	SEQUENCE 44 AA; 5240 MW; 04652A8DE5405D4F CRC64;				
Query Match	26.3%;	Score 25.5; DB 2; Length 44;			
Best Local Similarity	44.4%;	Pred. No. 1.3e+04;			
Matches	8; Conservative	2; Mismatches	5; Indels	3; Gaps	1;
Qy	9 IAFKIV---SOEPA 19	: : : :			
Db	29 LAKVINNRAQEPA 42	: : : :			
RESULT 201					
Q6XA02	PRELIMINARY;	PRT; 49 AA.			
AC	Q6XA02;				
DT	05-JUL-2004 (TREMBlrel. 27, Created)				
DT	05-JUL-2004 (TREMBlrel. 27, Last sequence update)				
DT	05-JUL-2004 (TREMBlrel. 27, Last annotation update)				
DE	Basic fibroblast growth factor receptor I (fragment).				
GN	Name=FGFR1;				
OS	Equus caballus (Horse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.				
OX	NCBI_TaxID=9796;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Takafuji V.A., Crisman M.V., Seat K.L., Sharova L.V., Ward D.L.,				
RA	Howard R.D.;				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY246707; AAP78692.1; -.				
DR	GO; GO:0004872; F:receptor activity; IEA.				
DR	InterPro; IPR011009; Kinase like.				
DR	InterPro; IPR000719; Prot.Kinase.				
DR	ProDom; PD000001; Prot_kinase; 1.				
KW	Receptor.				
FT	NON TER 1				
FT	NON TER 49				
SQ	SEQUENCE 49 AA; 5560 MW; E04614AF0B981820 CRC64;				
Query Match	26.3%;	Score 25.5; DB 2; Length 49;			
Best Local Similarity	29.4%;	Pred. No. 1.5e+04;			
Matches	5; Conservative	7; Mismatches	4; Indels	1; Gaps	1;
Qy	1 EPNHLSKIAPKIVSQE 17	: : : : : :			
Db	7 KPNRV-TKVAVKMLKSD 22	: : : : : :			
RESULT 202					
AAP78692	PRELIMINARY;	PRT; 49 AA.			
ID	AAP78692				
AC	AAP78692;				
DT	02-MAR-2004 (TREMBlrel. 27, Created)				
DT	02-MAR-2004 (TREMBlrel. 27, Last sequence update)				
DT	02-MAR-2004 (TREMBlrel. 27, Last annotation update)				
DE	Basic fibroblast growth factor receptor I (fragment).				
GN	FGFR1.				
OS	Equus caballus (Horse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.				
OX	NCBI_TaxID=9796;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Takafuji V.A., Crisman M.V., Seat K.L., Sharova L.V., Ward D.L.,				
RA	Howard R.D.;				
RT	"Expression analysis of equine interleukin-1b treated equine synovium				
RT	using suppression subtractive hybridization analysis (SSH-PCR).";				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY246707; AAP78692.1; -.				
KW	Receptor.				
FT	NON TER 1				
FT	NON TER 49				
SQ	SEQUENCE 49 AA; 5560 MW; E04614AF0B981820 CRC64;				
Query Match	26.3%;	Score 25.5; DB 2; Length 49;			
Best Local Similarity	29.4%;	Pred. No. 1.5e+04;			
Matches	5; Conservative	7; Mismatches	4; Indels	1; Gaps	1;
Qy	1 EPNHLSKIAPKIVSQE 17	: : : : : :			

D6
7 KPNRV-TKVAVKMLKSD 22

RESULT 203

ID	MIL17_BOVIN	STANDARD;	PRT;	20 AA.
AC	P35451;			
DT	01-JUN-1994	(Rel. 29, Created)		
DT	01-JUN-1994	(Rel. 29, Last sequence update)		
DT	03-JUL-2004	(Rel. 44, Last annotation update)		
DE	17 kDa milk glycoprotein (fragment).			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
[1]				
RN				
RP	SEQUENCE.			
RC	TISSUE=Milk;			
RX	MEDLINE=93308294; PubMed=8320368;			
RA	Soerensen E.S., Petersen T.E.;			
RT	"Purification and characterization of three proteins isolated from the			
RL	protease peptone fraction of bovine milk.";			
RL	J. Dairy Res. 60:189-197(1993).			
CC	-1- FTM: N-glycosylated.			
CC	-1- SIMILARITY: TO camel whey protein.			
KW	Direct protein sequencing; Glycoprotein; Milk.			
FT	NON_TER	1		
FT	NON_TER	20		
FT	NON_TER	20		
SQ	SEQUENCE	20 AA; 4CCA589404C62C27		CRG64;

Query Match	Score 25;	DB 1;	Length 20;
	25.8%		

Best Local Similarity 23.5%; Pred. No. 7e+03;
Matches 4; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```
QY      1 EPNHLNSKIAFKIVSQE 17
      :| | : | : :
DB      4 QPOSONPKLPLSLIKEK 20
```

RESULT 20A

AC	NRLA ACISA	STANDARD;	PRT;	21 AA.
ID	P33036;			
DT	01-OCT-1993	(Rel. 27, Created)		
DT	01-OCT-1993	(Rel. 27, Last sequence update)		
DT	05-JUL-2004	(Rel. 44, Last annotation update)		
DE	Nitrilase [EC 3.5.5.1] (Fragment).			
OS	Acinetobacter sp. (strain AK226).			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;			
OC	Moraxellaceae; Acinetobacter.			
ON	NCBI_TaxID=72569;			
RX	[1]			
RP	SEQUENCE			
RP	MEDLINE=91345837; PubMed=1369128;			
RA	Yamamoto K., Komatsu K.;			
RT	"Purification and characterization of nitrilase responsible for the			
RT	enantioselective hydrolysis from <i>Acinetobacter</i> sp. AK 226.";			
RL	<i>Agric. Biol. Chem.</i> 55:1459-1466(1991).			
CC	-!- FUNCTION: Acts on many kinds of nitrile compounds such as			
CC	aliphatic, aromatic, and heterocyclic mononitriles or dinitriles.			
CC	Prefers S-(-)-2-(4'-isobutylphenyl)-propionitrile to R-(+)-2-(4'-			
CC	isobutylphenyl)-propionitrile as the substrate.			
CC	-!- CATALYTIC ACTIVITY: A nitrile + H(2)O = a carboxylate + NH(3).			
CC	-!- SIMILARITY: Belongs to the CN hydrolase family. Nitrilase			
CC	subfamily.			
DR	InterPro: IPR000132; N/CN hydratase.			
DR	InterPro: IPR003010; Ntase/CNhydrtase.			
DR	PROSITE; PS50263; CN HYDROLASE; PARTIAL.			
DR	PROSITE; PS00920; NITRIL_CHT 1; PARTIAL.			
DR	PROSITE; PS00921; NITRIL_CHT 2; PARTIAL.			
KW	Direct protein sequencing; Hydrolase.			

FT	NON_TER	21	21
SO	SEQUENCE	21 AA:	222

Query Match 25.8%; Score 25; DB 1; Length 21;
Best Local Similarity 46.2%; Pred. No. 7.4e+03;
Matches 6; Conservative 0; Mismatches 7; Indels

QY	6	NSKIAFKIVSQEP	18
Db	4	NSKFLAATVQAE	16

RESULT 205

ID	Q79CG1	PRELIMINARY;	PRT;	21 AA.
AC	Q79CG1;			
DT	05-JUL-2004	(T-EMBLrel. 27, Created)		
DT	05-JUL-2004	(T-EMBLrel. 27, Last sequence update)		
DT	05-JUL-2004	(T-EMBLrel. 27, Last annotation update)		
DE	Virulence associated protein C (fragment) .			
GN	Name=vapC;			
OS	Bacteroides nodosus (Dichelobacter nodosus) .			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Cardiobacteriales;			
OC	Cardiobacteriaceae; Dichelobacter.			
OX	NCBI_TaxID=870;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=HA338;			
RC	MEDLINE=97071332; PubMed=8914257;			
RX				
RA	Rood J.I., Howarth P.A., Haring V., Billington S.J., Yong W.K.,			
RA	Liu D., Palmer M.A., Pittman D.R., Links I., Stewart D.J.,			
RA	Vaughan J.A.;			
RT	"Comparison of gene probe and conventional methods for the			
RT	differentiation of ovine footrot isolates of Dichelobacter nodosus." ;			
RL	Vet. Microbiol 52:127-141(1996).			
RL				
DR	ENBL; U39547; AAC44625.1; -.			
DR	NON TER			
FT	21			
FT	SEQUENCE 21 AA; 2477 MW; 5E28E5F03A316A68 CRC64;			

Query Match 25.8%; Score 25; DB 2; Length 21;
Best Local Similarity 28.6%; Pred.No. 7.4e+03;
Matches 4; Conservative 6; Mismatches 4; Indels

```
Qy      5 LNSKIAFKIVSQEP 18
         |::| |:::|
Db      7 LDTNICIYIINRKP 20
```

RESULT 206

AC	POQA_PSEST	STANDARD;	PRT;	23 AA.
ID	POQA_PSEST			
AD	Q9F9U1;			
DT	10-OCT-2003	(Rel. 42, Created)		
DT	10-OCT-2003	(Last sequence update)		
DT	05-JUL-2004	(Rel. 44, Last annotation update)		
DE	Coenzyme PQQ synthesis protein A (Pyrroloquinoline quinone biosynthesis protein A).			
DE	Name=pgqA;			
GN	Pseudomonas stutzeri (Pseudomonas perfectomarina).			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;			
OC	Pseudomonadaceae; Pseudomonas.			
OX	NCBI_TaxId=316;			
	[1]			
RN	SEQUENCE FROM N.A.			
RP	SEQUENCE=BC-2;			
RC	STRAIN=BC-2;			
RA	Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,			
RA	Terwilliger T.C.			
RT	"Identification and characterization of genes activated by 2-chloroethanol in Pseudomonas stutzeri BC-2.";			
RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.			
CC	!- FUNCTION: Required for coenzyme pyrroloquinoline quinone (PQQ) biosynthesis. Probably provides the glutamate and tyrosine residues that are cross-linked and modified to form the coenzyme.			
CC				
CC				
CC				

```

CC (By similarity).
CC -1- PATHWAY: Pyrroloquinoline quinone (PQQ) biosynthesis.
CC -1- SIMILARITY: Belongs to the pqqA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF176640; AAC09250.1; -.
CC HAMAP, MF_00656; -, 1.
CC PQQ; PQQ Biosynthesis.
CC CROSSLNK 15
CC FT
CC SEQUENCE 23 AA; 2809 MW; 80815999B7EBC47 CRC64;
CC -----
CC Query Match 25.8%; Score 25; DB 1; Length 23;
CC Best Local Similarity 23.1%; Pred. No. 8.1e+03;
CC Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC QY 1 EPNHLNSKIAPKI 13
CC DB 4 KPSYTDLRIGFEV 16
CC -----
CC RESULT 207
CC RR19 PETHY STANDARD; PRT; 23 AA.
CC AC P18550;
CC DT 01-NOV-1990 (Rel. 16, Created)
CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
CC DT 05-JUL-2004 (Rel. 44, Last annotation update)
CC DE Chloroplast 30S ribosomal protein S19 (Fragment).
CC GN Name=rrs19;
CC OS Petunia hybrida (Petunia).
CC OG Chloroplast.
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
CC OC lamiales; Solanales; Solanaceae; Petunia.
CC OX NCBI_TaxID=4102;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=89136067; PubMed=3224388;
CC RA Aldrich J., Cherney B.W., Williams C., Merlin E.;
CC RT "Sequence analysis of the junction of the large single copy region and
CC the large inverted repeat in the petunia chloroplast genome.";
CC RL Curr. Genet. 14:487-492(1988).
CC CC -1- FUNCTION: Protein S19 forms a complex with S13 that binds strongly
CC to the 16S ribosomal RNA (by similarity).
CC CC -1- SIMILARITY: Belongs to the S19p family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M37322; AAB02425.1; -.
CC HAMAP, MF_00531; -, 1.
CC InterPro: IPR002222; Ribosomal_S19.
CC DR Pfam: PF00203; Ribosomal_S19; 1.
CC DR PROSITE: PS00323; RIBOSOMAL_S19; PARTIAL.
CC KW Chloroplast; Ribosomal protein; rRNA-binding.
CC FT INIT MET 0 By similarity.
CC FT NON_TER 23 23
CC SEQUENCE 23 AA; 2665 MW; CA20D0128FC9AA49 CRC64;
CC -----
CC Query Match 25.8%; Score 25; DB 1; Length 23;

```

```

Best Local Similarity 71.4%; Pred. No. 8.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NHIINSKI 9
DB 12 NHIINSKI 18

RESULT 208
Q9S444 PRELIMINARY; PRT; 23 AA.
ID Q9S444;
AC Q9S444;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
CC NCBI_TaxID=287;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=R545;
CC RA Roy D., Coulombe M., Perron K., Roy P.H.;
CC RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC DR EMBL: AF162771; AAD46627.1; -.
CC KW Hypothetical protein.
CC SQ SEQUENCE 23 AA; 2236 MW; 1A3F2B45EE00D47D CRC64;

Query Match 25.8%; Score 25; DB 2; Length 23;
Best Local Similarity 66.7%; Pred. No. 8.1e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNHLNS 7
DB 17 PSHLSS 22

RESULT 209
Q99LY6 PRELIMINARY; PRT; 25 AA.
ID Q99LY6;
AC Q99LY6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P1k3r1 protein (Fragment).
GN Name=P1k3r1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=mx FVB/N;
CC RX TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
CC RX MEDLINE=22388257; PubMed=12477932;
CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
CC RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
CC RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
CC RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
CC RA Stapleton M., Soares M.B., Botaldo M.F., Casavant T.L., Scheetz T.E.,
CC RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
CC RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
CC RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
CC RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
CC RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
CC RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
CC RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
CC RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
CC RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
CC RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
CC RA Jones S.J., Marra M.A.;

```



```

RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP [2]
RC STRAIN=mix FVB/N;
RA TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002168; AA02168.1; -.
DR HSP; P23727; IBI.
DR MGI; 97583; Pk3r1.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0030183; P:B-cell differentiation; IMP.
FT NON TER 1
SQ SEQUENCE 25 AA; 2924 MW; 3F757937C4B2C95C CRC64;

Query Match 25.8%; Score 25; DB 2; Length 25;
Best Local Similarity 30.8%; Pred. No. 8.9e+03;
Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQE 17
Db 11 LNTVLAVPVYAAQQ 23

RESULT 210
Q9QX08 PRELIMINARY; PRT; 27 AA.
AC Q9QX08
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Nonmuscle tropomyosin 5 (Fragment).
GN Name=Tpm5;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98325069; PubMed=9660825;
RA Dufour C., Weinberger R.P., Schevzov G., Jeffrey P.L., Gunning P.;
RT "Splicing of two internal and four carboxyl-terminal alternative exons
RT in nonmuscle tropomyosin 5 pre-mRNA is independently regulated during
RT development.";
RL J. Biol. Chem. 273:18547-18555(1998).
DR EMBL; AF053359; AAC27290.1; -.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00261; Tropomyosin; 1.
FT NON TER 1
SQ SEQUENCE 27 AA; 3125 MW; EF4B323BDBF61543 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 27;
Best Local Similarity 40.0%; Pred. No. 9.6e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 KIAFKIVSQE 17
Db 7 KLYKKAISDE 16

RESULT 211
Q96J98 PRELIMINARY; PRT; 29 AA.
AC Q96J98
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Endothelial nitric oxide synthase (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Artery;
RA Zhang Y., Freedman J.E.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF387340; AAK71989.1; -.
FT NON TER 29
SQ SEQUENCE 29 AA; 2795 MW; 3F7F3F5E1973B4B CRC64;

Query Match 25.8%; Score 25; DB 2; Length 29;
Best Local Similarity 71.4%; Pred. No. 1e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 KIVSQEP 18
Db 5 KSVAQEP 11

RESULT 212
Q6JL84 PRELIMINARY; PRT; 29 AA.
AC Q6JL84
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FA19;
RX PubMed=15084227;
RA Snyder L.A., Davies J.K., Saunders N.J.;
RT "Microarray genotyping of key experimental strains of Neisseria
RT gonorrhoeae reveals gene complement diversity and five new neisserial
RT genes associated with Minimal Mobile Elements.";
RL BMC Genomics 5:23-23(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FA19;
RA Snyder L.A.S., Davies J.K., Saunders N.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY386270; AAS16535.1; -.
KW Hypothetical protein.
FT NON TER 29
SQ SEQUENCE 29 AA; 3348 MW; 71E842B01E42331D CRC64;

Query Match 25.8%; Score 25; DB 2; Length 29;
Best Local Similarity 50.0%; Pred. No. 1e+04;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNHLSN 7
Db 18 PNHHA 23

RESULT 213
AAS16535 PRELIMINARY; PRT; 29 AA.
AC AAS16535
DT 20-MAY-2004 (TrEMBLrel. 27, Created)
DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 20-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=FA19;
RX PubMed=15084227;
RA Snyder L.A., Davies J.K., Saunders N.J.;
RT "Microarray genotyping of key experimental strains of *Neisseria gonorrhoeae* reveals gene complement diversity and five new *neisseria* genes associated with Minimal Mobile Elements.";
RL BMC Genomics 5:23-23(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FA19;
RA Snyder L.A.S., Davies J.K., Saunders N.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY386270; AAS16535.1; -.
KW Hypothetical protein.
FT NON TER 29
SQ SEQUENCE 29 AA; 3348 MW; 71E842B01E42331D CRC64;

Query Match 25.8%; Score 25; DB 2; Length 29;
Best Local Similarity 50.0%; Pred. No. 1e+04;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNHLMS 7
DB 18 FNHIHA 23

RESULT 214
Q9R8A1 PRELIMINARY; PRT; 30 AA.
AC Q9R8A1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Beta-glucuronidase (Fragment).
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]_TaxID=813;
RP SEQUENCE FROM N.A.
RC STRAIN=L2 434B;
RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087267; AAD04045.1; -.
FT NON TER 1
FT NON TER 30
SQ SEQUENCE 30 AA; 3207 MW; A6494B849B994E29 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 30;
Best Local Similarity 41.7%; Pred. No. 1.1e+04;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 PNHLMSKIAFKI 13
DB 19 PDSLNTVVTGKI 30

RESULT 215
Q732G1 PRELIMINARY; PRT; 30 AA.
AC Q732G1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedDocuNames=BCB3953;
OS *Bacillus cereus* (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]_TaxID=222523;
RP SEQUENCE FROM N.A.
RX PubMed=14960714;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,

Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F., Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017276; AAS42856.1; -.
DR TIGR; BCE3953; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 30 AA; 3324 MW; 59C45ED531FE29FB CRC64;

Query Match 25.8%; Score 25; DB 2; Length 30;
Best Local Similarity 44.4%; Pred. No. 1.1e+04;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 LNSKIAFKI 13
DB 6 INSKVAIVV 14

RESULT 216
AAS42856 PRELIMINARY; PRT; 30 AA.
ID AAS42856
AC AAS42856;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN BCE3953.
OS *Bacillus cereus* (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L., Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F., Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017276; AAS42856.1; -.
DR TIGR; BCE3953; -.
KW Hypothetical protein.
SQ SEQUENCE 30 AA; 3324 MW; 59C45ED531FE29FB CRC64;

Query Match 25.8%; Score 25; DB 2; Length 30;
Best Local Similarity 44.4%; Pred. No. 1.1e+04;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 LNSKIAFKI 13
DB 6 INSKVAIVV 14

RESULT 217
O14937 PRELIMINARY; PRT; 31 AA.
ID O14937
AC O14937;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NOS III (Fragment).
GN Name=NOS3;
OS *Homo sapiens* (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Kleintert H., Koller K.P., Wohlfart P., Jensen T., Forstermann U.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF032908; AAB87083.1; -.

```
FT NON TER 31
SQ SEQUENCE 31 AA; 2949 MW; 555B7F7F3F35E197 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 31;
Best Local Similarity 71.4%; Pred. No. 1.1e+04;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 12 KIVSQEP 18
| : |||
Db 5 KVAQEP 11

RESULT 218
Q837M8 PRELIMINARY; PRT; 31 AA.
AC Q837M8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=EF0808;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacilliales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RC SEQUENCE FROM N.A.
RF STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Dougherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA Kouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis."
RL Science 299:2071-2074(2003).
DR EMBL; AE016949; AA080621.1; -.
DR TIGR; EF0808; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 31 AA; 3698 MW; E073A842FC5B7F9F CRC64;

Query Match 25.8%; Score 25; DB 2; Length 31;
Best Local Similarity 35.3%; Pred. No. 1.1e+04;
Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSOE 17
| : |||
Db 4 EKSYLEKSAFCFLKE 20

RESULT 219
Q81AC3 PRELIMINARY; PRT; 32 AA.
AC Q81AC3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Methyltransferase (EC 2.1.1.-).
GN ORFNames=BC3655;
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RC SEQUENCE FROM N.A.
RF MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapatal V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Hasekorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.C.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
```

```
RT Bacillus anthracis.";
RL Nature 423:87-91(2003).
DR EMBL; AE017009; AAP1084.1; -.
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Methyltransferase; Transferase.
SQ SEQUENCE 32 AA; 4059 MW; 5BB588029718CFC4 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 32;
Best Local Similarity 57.1%; Pred. No. 1.1e+04;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
| : |||
Db 12 NHLQTRI 18

RESULT 220
Q9QZQ2 PRELIMINARY; PRT; 32 AA.
AC Q9QZQ2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Neurotensin receptor (Fragment).
GN Name=Ntsr; Synonyms=Ntrl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RF STRAIN=129;
RX MEDLINE=994445567; PubMed=10514493;
RA Tavares D., Tully K., Dobner P.R.;
RT "Sequences required for induction of neurotensin receptor gene
RT expression during neuronal differentiation of N1E-115 neuroblastoma
RT cells.";
RL J. Biol. Chem. 274:30066-30079(1999).
DR EMBL; AF172326; AAD51806.1; -.
DR MGD; MGI:97386; Ntsr.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON TER 32
SQ SEQUENCE 32 AA; 3447 MW; 7F7EA4FA2CFC2EFB CRC64;

Query Match 25.8%; Score 25; DB 2; Length 32;
Best Local Similarity 66.7%; Pred. No. 1.1e+04;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 HLNSKI 9
| : |||
Db 2 HLNSV 7

RESULT 221
Q64854 PRELIMINARY; PRT; 32 AA.
AC Q64854;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNA-binding protein (Fragment).
GN Name=DBP;
OS Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
OX NCBI_TaxID=10553;
RN [1]
RC SEQUENCE FROM N.A.
RF MEDLINE=93362429; PubMed=8395124;
RA Cai F., Weber J.M.;
RT "Organization of the avian adenovirus genome and the structure of its
RT endopeptidase.";
```

RL Virology 196:358-362(1993).
 DR EMBL; L13161; ARA51403.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 KW DNA-binding.
 FT NON TER
 SQ SEQUENCE 32 AA; 3584 MW; F6841332592BE037 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 32;
 Best Local Similarity 35.7%; Pred. No. 1.1e+04;
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIATKIVSQEP 18
 || : : : :
 Db 7 LNPFSVLNKASEEP 20

RESULT 222

Q9YFG7 PRELIMINARY; PRT; 32 AA.
 ID Q9YFG7
 AC Q9YFG7
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Nonstructural polyprotein (Fragment).
 OS Highlands J virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 CC Alphavirus; VEEV complex.
 OX NCBI_TaxID=11024;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83010261; PubMed=62888962;
 RA Ou J.H., Trent D.W., Strauss J.H.;
 RT "The 3'-non-coding regions of alphavirus RNAs contain repeating
 sequences."; J. Mol. Biol. 156:719-730(1982).
 RL J. Mol. Biol. 156:719-730(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83268700; PubMed=6308269;
 RA Ou J.H., Strauss E.G., Strauss J.H.;
 RT "The 5'-terminal sequences of the genomic RNAs of several
 alphaviruses."; J. Mol. Biol. 168:1-15(1983).
 RL J. Mol. Biol. 168:1-15(1983).
 DR EMBL; K00700; AAA45382.1; -.
 KW Polyprotein.
 FT NON TER
 SQ SEQUENCE 32 AA; 3577 MW; 2084B13FED847058 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 32;
 Best Local Similarity 31.6%; Pred. No. 1.1e+04;
 Matches 6; Conservative 4; Mismatches 5; Indels 4; Gaps 1;

QY 3 NMLN----SKIAFKIVSQE 17
 :|| :|| :|| :|| :|| :||
 Db 6 DHANARAFSHVATKLIIE 24

RESULT 223

Q7LIB6 PRELIMINARY; PRT; 33 AA.
 ID Q7LIB6
 AC Q7LIB6
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Gmp synthase (EC 6.3.5.2) (Fragment).
 OS Lactobacillus delbrueckii (subsp. lactis).
 CC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 OX NCBI_TaxID=29397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 4797;
 RA Langenheilm J.F., Ulrich R.L.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBDJ databases.
 DR EMBL; AF496339; AAQ07021.1; -.
 DR GO; GO:0003922; F:GMP synthase (glutamine-hydrolyzing) activity; IEA.
 DR GO; GO:0016874; F:Ligase activity; IEA.
 KW Ligase.
 FT NON TER
 SQ SEQUENCE 33 AA; 3830 MW; 6292C1C3308C8C46 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 33;
 Best Local Similarity 23.1%; Pred. No. 1.2e+04;
 Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 7 SKIAFKIVSQEPA 19
 : : : : : :
 Db 16 NRVLVDYVTSKPPA 28

RESULT 224

O88440 PRELIMINARY; PRT; 33 AA.
 ID O88440
 AC O88440
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Nonmuscle tropomyosin 5 (Fragment).
 GN Name=Tpm5;
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=98325069; PubMed=9660825;
 RA Dufour C., Weinberger R.P., Schvezov G., Jeffrey P.L., Gunning P.;
 RT "Splicing of two internal and four carboxyl-terminal alternative exons
 in nonmuscle tropomyosin 5 pre-mRNA is independently regulated during
 development."; J. Biol. Chem. 273:18547-18555(1998).
 RL J. Biol. Chem. 273:18547-18555(1998).
 DR EMBL; AF053360; AAC27291.1; -.
 DR InterPro; IPR000533; Tropomyosin.
 DR Pfam; PF00261; Tropomyosin; 1.
 FT NON TER
 SQ SEQUENCE 33 AA; 3723 MW; 5FF5861B89A38F4B CRC64;

Query Match 25.8%; Score 25; DB 2; Length 33;
 Best Local Similarity 40.0%; Pred. No. 1.2e+04;
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 KIAFKIVSQE 17
 : : : : : :
 Db 7 KLKYKAISDE 16

RESULT 225

AAQ07021 PRELIMINARY; PRT; 33 AA.
 ID AAQ07021
 AC AAQ07021
 DT 02-MAR-2004 (TREMBLrel. 27, Created)
 DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
 DE Gmp synthase (EC 6.3.5.2) (Fragment).
 OS Lactobacillus delbrueckii (subsp. lactis).
 CC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 OX NCBI_TaxID=29397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 4797;
 RA Langenheilm J.F., Ulrich R.L.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBDJ databases.
 DR EMBL; AF496339; AAQ07021.1; -.
 KW Ligase.

```

FT NON TER 1 1
SQ SEQUENCE 33 AA; 3830 MW; 6292C1C3308C8C46 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 33;
Best Local Similarity 23.1%; Pred. No. 1.2e+04;
Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SKTAFKIVSQEPA 19
   :::::|:|
Db 16 NRVDYDVTSKPFA 28

RESULT 226
RT36 BOVIN
ID RT36_BOVIN STANDARD; PRT; 34 AA.
AC P82908;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Mitochondrial 28S ribosomal protein S36 (S36mt) (MRP-S36) (Fragments).
GN Name=MRP36;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=21276436; PubMed=11279123; DOI=10.1074/jbc.M100727200;
RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;
RT "The small subunit of the mammalian mitochondrial ribosome:
RT identification of the full complement of ribosomal proteins present.";
RL J. Biol. Chem. 276:19363-19374(2001).
CC -!- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
KW Direct protein sequencing; Mitochondrion; Ribosomal protein.
FT NON TER 1 1
FT NON_CONS 18 19
FT NON TER 34 34
SQ SEQUENCE 34 AA; 3866 MW; 7CDA9FA4B08A412A CRC64;

Query Match 25.8%; Score 25; DB 1; Length 34;
Best Local Similarity 83.3%; Pred. No. 1.2e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 12 KIVSQE 17
   :|:|
Db 18 KIVSQE 23

RESULT 227
Q85NZ6
ID Q85NZ6 PRELIMINARY; PRT; 34 AA.
AC Q85NZ6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase subunit 1-like protein (Fragment).
GN Name=ND1;
OS Thecabius populi-monis.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
OC Aphidoidea; Pemphigidae; Pemphiginae; Thecabiinae;
OX NCBI_TaxID=220098;
RN [1]
RP SEQUENCE FROM N.A.
RA Abbot P., Withgott J.H.;
RT "Phylogenetic and molecular evidence for allochronic speciation in
RT gall-forming aphids (Pemphigus).";
RL Evolution 58:539-553(2004).

DR EMBL; AY182360; AAC47640.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER 1 1
FT NON TER 34 34
SQ SEQUENCE 34 AA; 4106 MW; 0EE479AD380D5670 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 34;
Best Local Similarity 50.0%; Pred. No. 1.2e+04;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PNLHLSKIAF 11
   ||::||
Db 25 PNNMNPFIFF 34

RESULT 228
Q73KJ5
ID Q73KJ5 PRELIMINARY; PRT; 34 AA.
AC Q73KJ5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=TDE2223;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidse T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
RA Gregeorgis E., Geer K., Tsegaye G., Malek J.A., Ayodeji B.,
RA Shatsman S., McLeod M.P., Smajic D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstein G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
RL EMBL; AE017253; AAS12742.1; -.
RW TIGR; TDE2223; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 34 AA; 3684 MW; E68C4F1CA95A305D CRC64;

Query Match 25.8%; Score 25; DB 2; Length 34;
Best Local Similarity 33.3%; Pred. No. 1.2e+04;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFK 12
   :|:|:|
Db 9 QPHCLSAEFAFQ 20

RESULT 229
Q87KH8
ID Q87KH8 PRELIMINARY; PRT; 34 AA.
AC Q87KH8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VP2998.
GN OrderedLocusNames=VP2998;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
SQ STRAIN=RIMD 2210633 / Serotype O3:K6;

```

```

RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749 (2003).
DR EMBL; AP005083; BAC61262.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 34 AA; 3840 MW; 011C95FB08EF9BF9 CRC64;

Query Match      25.8%; Score 25; DB 2; Length 34;
Best Local Similarity 36.4%; Pred. No. 1.2e+04;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 HNSKIAPKIV 14
Db 3 HVNLPITFVLI 13

RESULT 230
Q8CLO6 PRELIMINARY; PRT; 34 AA.
AC Q8CLO6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical
GN OrderedNames=y0581;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137963; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611 (2002).
DR EMBL; AE013660; AAM84169.1; -.
KW Hypothetical protein.
SQ SEQUENCE 34 AA; 3778 MW; FD69650811E489B3 CRC64;

Query Match      25.8%; Score 25; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNHL 5
Db 8 PNHL 11

RESULT 231
O70225 PRELIMINARY; PRT; 34 AA.
AC O70225;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE MARR1C6A (Fragment).
DE MARR1C6A (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brain;
RX Liao B.S., Jin W.L., Ju G.;

```

```

RL Submitted (JUN-1997) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF010441; AAC4890.1; -.
FT NON_TER 1
SQ SEQUENCE 34 AA; 3720 MW; 07995C33C5B599E7 CRC64;

Query Match      25.8%; Score 25; DB 2; Length 34;
Best Local Similarity 38.5%; Pred. No. 1.2e+04;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQEP 18
Db 5 SSHPAYKPIAQAP 17

RESULT 232
Q9ET72 PRELIMINARY; PRT; 34 AA.
AC Q9ET72;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Na/Ca exchanger (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=Liver;
RX MEDLINE=20358717; PubMed=10902911;
RA Wakimoto K., Kuro-O M., Yanaka N., Omori K., Komuro I., Imai Y.,
RA Nabeshima Y.;
RT "Isolation and characterization of Na(+) /Ca(2+) exchanger gene and
RT splicing isoforms in mice.";
RL DNA Seq. 11:75-81 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=Liver;
RA Koji W.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB010887; BAB08171.1; -.
FT NON_TER 1
FT NON_TER 34
SQ SEQUENCE 34 AA; 4226 MW; D7A379E0357815C1 CRC64;

Query Match      25.8%; Score 25; DB 2; Length 34;
Best Local Similarity 44.4%; Pred. No. 1.2e+04;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 AFKIVSQEP 18
Db 18 SFSVLVLEP 26

RESULT 233
AAS12742 PRELIMINARY; PRT; 34 AA.
ID AAS12742;
AC AAS12742;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN TDE2223.
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J., Durkin S.A.,

```

```

RA Daugherty S.C., Shetty J., Shvartsbeyn A., Gebregorgis E., Geer K.,
RA Tsagave G., Malek J.A., Ayodeji B., Shatsman S., McLeod M.P.,
RA Smajs D., Howell J.K., Pal S., Amin A., Vashish P., McNeill T.Z.,
RA Xiang Q., Sodergren E., Baca E., Weinstein G.M., Norris S.J.,
RA Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT with other spirochete genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL; AE017253; AAS12742.1; -.
DR TIGR; TDE2223; -.
KW Hypothetical protein.
SQ SEQUENCE 34 AA; 3684 MW; B68C4F1CA95A305D CRC64;
    Query Match      25.8%; Score 25; DB 2; Length 34;
    Best Local Similarity 33.3%; Pred. No. 1.2e+04;
    Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLNSKIAPK12
   : : : : :
Db 9 QPHCLSAEFAFQ 20
   : : : : :

RESULT 234
Q6QRN3 PRELIMINARY; PRT; 35 AA.
AC Q6QRN3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hydroxysteroid dehydrogenase-like protein (Fragment).
GN Name=HDLF;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Shin M.L., Cui X.S., Park S.Y., Kim E.Y., Park S.P., Lee W.J.,
RA Hwang K.C., Kim N.H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY528253; AAS20601.1; -.
FT NON TER 1
SQ SEQUENCE 35 AA; 3923 MW; 4FCA764BD2C4642C CRC64;
    Query Match      25.8%; Score 25; DB 2; Length 35;
    Best Local Similarity 57.1%; Pred. No. 1.3e+04;
    Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NHLNSKI 9
   : : : : :
Db 29 NQWNSKL 35
   : : : : :

RESULT 235
Q701K0 PRELIMINARY; PRT; 35 AA.
AC Q701K0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcriptional regulator FruR (Fragment).
GN Name=FruR;
OS Spiroplasma sp. NSRO.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Spiroplasmataceae; Spiroplasma.
OX NCBI_TaxID=204397;
RN [1]
RP SEQUENCE FROM N.A.
RA Montenegro H., Solferini V.N., Klaczko L.B., Hurst G.D.D.;
RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ628445; CAF32405.1; -.

```

```

FT NON TER 1
SQ SEQUENCE 35 AA; 3850 MW; 9441A4DBC647EB8A CRC64;
    Query Match      25.8%; Score 25; DB 2; Length 35;
    Best Local Similarity 31.2%; Pred. No. 1.3e+04;
    Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 EPNHLNSKIAPK1VQ 16
   : : : : :
Db 12 DKSKFNKSSVKFASQ 27
   : : : : :

RESULT 236
Q701K1 PRELIMINARY; PRT; 35 AA.
AC Q701K1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcriptional regulator FruR (Fragment).
GN Name=FruR;
OS Spiroplasma sp. MSRO.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Spiroplasmataceae; Spiroplasma.
OX NCBI_TaxID=265049;
RN [1]
RP SEQUENCE FROM N.A.
RA Montenegro H., Solferini V.N., Klaczko L.B., Hurst G.D.D.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ628444; CAF32403.1; -.
FT NON TER 1
SQ SEQUENCE 35 AA; 3850 MW; 9441A4DBC647EB8A CRC64;
    Query Match      25.8%; Score 25; DB 2; Length 35;
    Best Local Similarity 31.2%; Pred. No. 1.3e+04;
    Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 EPNHLNSKIAPK1VQ 16
   : : : : :
Db 12 DKSKFNKSSVKFASQ 27
   : : : : :

RESULT 237
Q8E9Z1 PRELIMINARY; PRT; 35 AA.
AC Q8E9Z1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SO4119.
GN OrderedLocustNames=SO4119;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Imbraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Nealsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015843; AAN57092.1; -.

```


RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
 RA Xu J.-G., Zhao G.-P.;
 RT "Unique physiological and pathogenic features of Leptospira
 RT interrogans revealed by whole-genome sequencing.";
 RL Nature 422:888-893(2003).
 DR EMBL; AF011498; AAN50593.1; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 36 AA; 4211 MW; 7BE4FD1830510CD9 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 36;
 Best Local Similarity 50.0%; Pred. No. 1.3e+04;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 KIAFKIVSQE 17
 |||||:
 Db 10 KIHFKVSSD 19

RESULT 245
 Q9PC23 PRELIMINARY; PRT; 36 AA.
 AC Q9PC23
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE Hypothetical protein.
 GN OrderedLocusNames=Xf1962;
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.D.,
 RA Marques M.V., Martins E.A.B., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C. de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshakho M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 DR EMBL; AF004015; AAF84764.1; -.
 DR PIR; D82617; D82617.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 36 AA; 4067 MW; 77E4BADI752822E9 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 36;
 Best Local Similarity 33.3%; Pred. No. 1.3e+04;
 Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PNHLNSKIAFKIVSQ 16

Db 20 PQHLADKTSLYLKSK 34
 |||||:
 RESULT 246
 Q8U1K4 PRELIMINARY; PRT; 37 AA.
 ID Q8U1K4
 AC Q8U1K4
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein PF1202.
 GN OrderedLocusNames=PF1202;
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 3638;
 RX MEDLINE=21079003; PubMed=11210495;
 RA Robb F.T., Maeder D.L., Brown J.R., DiRuggiero J., Stump M.D.,
 RA Yeh R.K., Weiss R.B., Dunn D.M.;
 RT "Genomic sequence of hyperthermophile, Pyrococcus furiosus:
 RT implications for physiology and enzymology.";
 RL Meth. Enzymol. 330:134-157(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE010228; AAL81326.1; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 37 AA; 4351 MW; A0060EDD9D29E5C CRC64;

Query Match 25.8%; Score 25; DB 2; Length 37;
 Best Local Similarity 45.5%; Pred. No. 1.3e+04;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQ 16
 |||||:
 Db 15 NQKIFMRIVDE 25

RESULT 247
 Q7REN2 PRELIMINARY; PRT; 37 AA.
 ID Q7REN2
 AC Q7REN2
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=PY06108;
 OS Plasmodium yoelii yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=73239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17XNL;
 RX PubMed=12368865;
 RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteza M.,
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Sengen J.D., Koo H.L.,
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
 RA Shallow S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
 RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
 RA Carucci D.J.;
 RT "Genome sequence and comparative analysis of the model rodent malaria
 RT parasite Plasmodium yoelii yoelii.";

```

RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data. EAA18260.1; -.
DR EMBL; AALJ01002028; EAA18260.1; -.
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 4231 MW; D5D7185EB1204B24 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 37;
Best Local Similarity 18.8%; Pred. No. 1.3e+04;
Matches 3; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 PNHLNSKIAFKIVSQE 17
Db |||:::
22 PNFFIKNLSFDVILEQ 37

RESULT 248
QY5212 PRELIMINARY; PRT; 37 AA.
AC Q95Z12;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ceratotoxin 3 precursor (Fragment).
GN Names=Crt3;
OS Ceratitidis rosa (Natal fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Tephritoidea; Tephritidae; Ceratitidis.
OX NCBI_TaxID=56958;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22599974; PubMed=12714984;
RA Rosetto M., Marchini D., de Filippis T., Ciolfi S., Prati F.,
RA Quilici S., Dallai R.;
RT "The ceratotoxin gene family in the medfly Ceratitidis capitata and the
RT Natal fruit fly Ceratitidis rosa (Diptera: Tephritidae).";
RL Hereditiy 90:382-389(2003).
DR EMBL; AJ313431; CAC43788.1; -.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 17 Potential.
FT CHAIN 30 >37 ceratotoxin 3.
FT NON_TER 37 37
SQ SEQUENCE 37 AA; 3984 MW; CF4723949AD6BF75 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 37;
Best Local Similarity 60.0%; Pred. No. 1.3e+04;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 IAFKIVSQEP 18
Db |||:::
10 IAFTCVVAEP 19

RESULT 249
QY5239 PRELIMINARY; PRT; 37 AA.
AC Q55239;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Flavin-associated protein A (Fragment).
GN Name=fprA;
OS Synechocystis sp.
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1143;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PCC 6803;
RX MEDLINE=95035996; PubMed=7524725;
RA Sakamoto T., Los D., Higashi S., Wada H., Nishida I., Ohmori M.,

```

```

RA Murata N.;
RT "Cloning of omega 3 desaturase from cyanobacteria and its use in
RT altering the degree of membrane-lipid unsaturation.";
RL Plant Mol. Biol. 26:249-263(1994).
DR EMBL; D13780; BAA02923.1; -.
DR GO; GO:0004733; P:Pyridoxamine-phosphate oxidase activity; IEA.
DR GO; GO:0008615; P:Pyridoxine biosynthesis; IEA.
DR InterPro; IPR009002; FMN binding.
DR InterPro; IPR000659; Pyridox oxidase.
DR ProDom; PD006312; Pyridox oxidase; 1.
DR PROSITE; PS01064; PYRIDOX_OXIDASE; 1.
FT NON_TER 1 1
SQ SEQUENCE 37 AA; 4630 MW; 13C7AEA0672DF852 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 37;
Best Local Similarity 23.1%; Pred. No. 1.3e+04;
Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 PNHLNSKIAFKIV 14
Db |||:::
14 PSLHDLRLQFNLL 26

RESULT 250
QY8FIK6 PRELIMINARY; PRT; 37 AA.
AC Q8FIK6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=LA3129;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang J.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang Y.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AS011475; AAN50327.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 37 AA; 4513 MW; 89DF0AA197540952 CRC64;

Query Match 25.8%; Score 25; DB 2; Length 37;
Best Local Similarity 41.7%; Pred. No. 1.3e+04;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQE 17
Db |||:::
10 NSYLFNVISVE 21

Search completed: October 19, 2004, 19:13:49
Job time : 222 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 19, 2004, 19:05:08 ; Search time 37 Seconds
(without alignments)
34.055 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 85066

Minimum DB seq length: 19

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	39.2	22	4	US-08-475-955-55
2	33	34.0	45	4	US-08-270-767-56675
3	33	34.0	49	4	US-08-270-767-38499
4	33	34.0	49	4	US-08-270-767-53716
5	32	33.0	26	1	US-08-427-072-10
6	31	32.0	33	3	US-08-172-841-29
7	31	32.0	33	4	US-08-951-621-29
8	31	32.0	34	4	US-08-690-454-72
9	31	32.0	46	4	US-08-270-767-36721
10	31	32.0	46	4	US-08-288-143-87
11	31	32.0	48	4	US-08-497A-445
12	30	30.9	20	4	US-08-576-824A-445
13	30	30.9	20	4	US-08-205-258-898
14	30	30.9	27	4	US-08-205-258-899
15	30	30.9	33	3	US-09-101-146-15
16	30	30.9	33	3	US-09-194-296-13
17	30	30.9	39	4	US-08-149-476-545
18	30	30.9	39	4	US-08-821-619-381
19	30	30.9	40	4	US-08-340-812-3
20	29	29.9	28	2	US-07-829-462-3
21	29	29.9	31	1	US-08-459-064B-3
22	29	29.9	31	1	US-08-460-421A-3
23	29	29.9	31	5	FCT-US93-00909-3
24	29	29.9	34	3	US-09-036-315-22
25	29	29.9	34	4	US-09-515-965A-1907
26	29	29.9	34	4	US-09-515-965A-1907
27	29	29.9	34	4	US-09-515-965A-1907
28	29	29.9	34	4	US-09-515-965A-1907
29	29	29.9	34	4	US-09-515-965A-1907
30	29	29.9	34	4	US-09-515-965A-1907
31	29	29.9	34	4	US-09-515-965A-1907
32	29	29.9	34	4	US-09-515-965A-1907
33	29	29.9	34	4	US-09-515-965A-1907
34	29	29.9	34	4	US-09-515-965A-1907
35	29	29.9	34	4	US-09-515-965A-1907
36	29	29.9	34	4	US-09-515-965A-1907
37	29	29.9	34	4	US-09-515-965A-1907
38	29	29.9	34	4	US-09-515-965A-1907
39	29	29.9	34	4	US-09-515-965A-1907
40	29	29.9	34	4	US-09-515-965A-1907
41	28.5	29.4	49	4	US-08-812-008-12
42	28.5	29.4	49	4	US-08-812-008-12
43	28	28.9	22	4	US-08-851-567B-70
44	28	28.9	26	3	US-09-367-065A-12
45	28	28.9	26	3	US-09-367-065A-12
46	28	28.9	26	4	US-09-568-466-12
47	28	28.9	26	4	US-09-568-466-12
48	28	28.9	26	4	US-09-568-466-12
49	28	28.9	30	3	US-09-308-935-6
50	28	28.9	36	2	US-08-436-420-37
51	28	28.9	38	3	US-08-905-223-285
52	28	28.9	41	4	US-09-270-767-61791
53	28	28.9	42	4	US-03-205-258-557
54	28	28.9	42	4	US-09-270-767-35410
55	28	28.9	42	4	US-09-270-767-50627
56	28	28.9	43	4	US-09-106-568E-40
57	28	28.9	44	4	US-09-270-767-37531
58	28	28.9	44	4	US-09-270-767-52748
59	27	27.8	19	4	US-08-442-989-50
60	27	27.8	20	2	US-08-934-915-109
61	27	27.8	21	4	US-08-924-629C-12
62	27	27.8	21	4	US-08-924-629C-80
63	27	27.8	22	4	US-09-228-990-56
64	27	27.8	22	1	US-08-447-411-78
65	27	27.8	23	1	US-08-447-411-78
66	27	27.8	23	2	US-08-662-227-26
67	27	27.8	23	2	US-08-662-227-28
68	27	27.8	23	3	US-09-017-947-26
69	27	27.8	23	3	US-09-017-947-28
70	27	27.8	23	4	US-09-228-990-57
71	27	27.8	23	4	US-09-925-442-26
72	27	27.8	23	4	US-09-925-442-28
73	27	27.8	24	4	US-09-106-568B-111
74	27	27.8	24	4	US-09-228-990-58
75	27	27.8	24	4	US-09-270-767-57210
76	27	27.8	24	4	US-09-270-767-58528
77	27	27.8	25	4	US-09-228-990-59
78	27	27.8	26	3	US-09-480-993-3
79	27	27.8	26	3	US-09-367-065A-3
80	27	27.8	26	3	US-09-367-065A-10
81	27	27.8	26	3	US-09-367-065A-11
82	27	27.8	26	3	US-09-367-065A-14
83	27	27.8	26	3	US-09-367-065A-19
84	27	27.8	26	3	US-09-367-065A-20
85	27	27.8	26	4	US-09-228-990-60
86	27	27.8	26	4	US-09-568-466-3
87	27	27.8	26	4	US-09-568-466-10
88	27	27.8	26	4	US-09-568-466-11
89	27	27.8	26	4	US-09-568-466-14
90	27	27.8	26	4	US-09-568-466-19
91	27	27.8	26	4	US-09-568-466-20
92	27	27.8	27	4	US-08-924-629C-11
93	27	27.8	27	4	US-09-228-990-55
94	27	27.8	27	4	US-09-228-990-61
95	27	27.8	28	4	US-09-228-990-54
96	27	27.8	28	4	US-09-228-990-62
97	27	27.8	28	4	US-09-228-990-65
98	27	27.8	28	4	US-09-228-990-79
99	27	27.8	28	4	US-09-442-989-22
100	27	27.8	28	4	US-09-442-989-25

Sequence 1907, Ap

101	27	27.8	29	4	US-09-228-990-53	Sequence 53, Appl	174	27	27.8	41	2	US-08-392-625-25	Sequence 25, Appl
102	27	27.8	29	4	US-09-228-990-63	Sequence 63, Appl	175	27	27.8	41	2	US-08-456-961A-25	Sequence 25, Appl
103	27	27.8	30	1	US-08-248-505-4	Sequence 4, Appl	176	27	27.8	41	3	US-08-936-165A-290	Sequence 290, Appl
104	27	27.8	30	4	US-09-228-990-52	Sequence 52, Appl	177	27	27.8	47	4	US-10-080-505-18	Sequence 18, Appl
105	27	27.8	30	4	US-09-228-990-64	Sequence 64, Appl	178	27	27.8	49	3	US-08-890-865A-21	Sequence 21, Appl
106	27	27.8	31	4	US-09-228-990-3	Sequence 3, Appl	179	26.5	27.3	36	2	US-08-867-087B-59	Sequence 59, Appl
107	27	27.8	31	4	US-09-228-990-4	Sequence 4, Appl	180	26.5	27.3	50	3	US-09-177-249-167	Sequence 167, Appl
108	27	27.8	31	4	US-09-228-990-5	Sequence 5, Appl	181	26	26.8	19	3	US-09-308-935-15	Sequence 15, Appl
109	27	27.8	31	4	US-09-228-990-6	Sequence 6, Appl	182	26	26.8	20	2	US-08-564-972-52	Sequence 52, Appl
110	27	27.8	31	4	US-09-228-990-7	Sequence 7, Appl	183	26	26.8	21	2	US-08-310-912A-76	Sequence 76, Appl
111	27	27.8	31	4	US-09-228-990-8	Sequence 8, Appl	184	26	26.8	21	3	US-09-128-155-12	Sequence 12, Appl
112	27	27.8	31	4	US-09-228-990-9	Sequence 9, Appl	185	26	26.8	21	3	US-08-841-089-76	Sequence 76, Appl
113	27	27.8	31	4	US-09-228-990-10	Sequence 10, Appl	186	26	26.8	21	3	US-09-301-085-76	Sequence 76, Appl
114	27	27.8	31	4	US-09-228-990-11	Sequence 11, Appl	187	26	26.8	21	3	US-09-230-548-15	Sequence 15, Appl
115	27	27.8	31	4	US-09-228-990-12	Sequence 12, Appl	188	26	26.8	21	3	US-08-924-629C-78	Sequence 78, Appl
116	27	27.8	31	4	US-09-228-990-13	Sequence 13, Appl	189	26	26.8	21	4	US-08-924-629C-79	Sequence 79, Appl
117	27	27.8	31	4	US-09-228-990-14	Sequence 14, Appl	190	26	26.8	21	5	PCT-US95-04570-76	Sequence 76, Appl
118	27	27.8	31	4	US-09-228-990-15	Sequence 15, Appl	191	26	26.8	21	5	PCT-US95-04589-76	Sequence 76, Appl
119	27	27.8	31	4	US-09-228-990-20	Sequence 20, Appl	192	26	26.8	22	3	US-08-847-844A-55	Sequence 55, Appl
120	27	27.8	31	4	US-09-228-990-21	Sequence 21, Appl	193	26	26.8	24	4	US-09-690-454-130	Sequence 130, Appl
121	27	27.8	31	4	US-09-228-990-22	Sequence 22, Appl	194	26	26.8	24	4	US-09-843-221A-133	Sequence 133, Appl
122	27	27.8	31	4	US-09-228-990-23	Sequence 23, Appl	195	26	26.8	25	1	US-08-145-708A-20	Sequence 20, Appl
123	27	27.8	31	4	US-09-228-990-24	Sequence 24, Appl	196	26	26.8	25	2	US-08-331-454-20	Sequence 20, Appl
124	27	27.8	31	4	US-09-228-990-25	Sequence 25, Appl	197	26	26.8	26	3	US-09-230-548-26	Sequence 26, Appl
125	27	27.8	31	4	US-09-228-990-26	Sequence 26, Appl	198	26	26.8	26	3	US-09-843-221A-97	Sequence 97, Appl
126	27	27.8	31	4	US-09-228-990-27	Sequence 27, Appl	199	26	26.8	31	1	US-08-248-021A-5	Sequence 5, Appl
127	27	27.8	31	4	US-09-228-990-28	Sequence 28, Appl	200	26	26.8	31	2	US-08-031-538-49	Sequence 49, Appl
128	27	27.8	31	4	US-09-228-990-29	Sequence 29, Appl	201	26	26.8	34	4	US-09-843-221A-92	Sequence 92, Appl
129	27	27.8	31	4	US-09-228-990-30	Sequence 30, Appl	202	26	26.8	34	4	US-09-843-221A-128	Sequence 128, Appl
130	27	27.8	31	4	US-09-228-990-31	Sequence 31, Appl	203	26	26.8	35	3	US-08-995-369-7	Sequence 7, Appl
131	27	27.8	31	4	US-09-228-990-36	Sequence 36, Appl	204	26	26.8	35	5	PCT-US95-10075-7	Sequence 7, Appl
132	27	27.8	31	4	US-09-228-990-37	Sequence 37, Appl	205	26	26.8	36	3	US-08-651-136C-28	Sequence 28, Appl
133	27	27.8	31	4	US-09-228-990-38	Sequence 38, Appl	206	26	26.8	36	3	US-09-227-357-550	Sequence 550, Appl
134	27	27.8	31	4	US-09-228-990-39	Sequence 39, Appl	207	26	26.8	36	3	US-09-229-911A-28	Sequence 28, Appl
135	27	27.8	31	4	US-09-228-990-40	Sequence 40, Appl	208	26	26.8	37	3	US-09-230-548-31	Sequence 31, Appl
136	27	27.8	31	4	US-09-228-990-41	Sequence 41, Appl	209	26	26.8	38	1	US-08-176-500-79	Sequence 79, Appl
137	27	27.8	31	4	US-09-228-990-50	Sequence 50, Appl	210	26	26.8	38	1	US-08-471-052A-79	Sequence 79, Appl
138	27	27.8	31	4	US-09-228-990-74	Sequence 74, Appl	211	26	26.8	38	1	US-08-189-331-79	Sequence 79, Appl
139	27	27.8	31	4	US-09-228-990-81	Sequence 81, Appl	212	26	26.8	38	2	US-08-471-939-79	Sequence 79, Appl
140	27	27.8	31	4	US-09-228-990-82	Sequence 82, Appl	213	26	26.8	38	2	US-08-471-800-79	Sequence 79, Appl
141	27	27.8	31	4	US-09-228-990-83	Sequence 83, Appl	214	26	26.8	38	2	US-08-471-068-79	Sequence 79, Appl
142	27	27.8	31	4	US-09-228-990-84	Sequence 84, Appl	215	26	26.8	39	4	US-09-270-767-61935	Sequence 61935, A
143	27	27.8	31	4	US-09-228-990-85	Sequence 85, Appl	216	26	26.8	42	1	US-08-099-354-8	Sequence 2, Appl
144	27	27.8	31	4	US-09-228-990-86	Sequence 86, Appl	217	26	26.8	42	2	US-08-288-059-34	Sequence 34, Appl
145	27	27.8	31	4	US-09-228-990-87	Sequence 87, Appl	218	26	26.8	42	2	US-09-270-767-60941	Sequence 60941, A
146	27	27.8	31	4	US-09-228-990-88	Sequence 88, Appl	219	26	26.8	42	4	US-09-270-767-32405	Sequence 32405, A
147	27	27.8	31	4	US-09-442-989-1	Sequence 1, Appl	220	26	26.8	44	4	US-09-270-767-47622	Sequence 47622, A
148	27	27.8	31	4	US-09-442-989-2	Sequence 2, Appl	221	26	26.8	46	3	US-09-100-600A-43	Sequence 43, Appl
149	27	27.8	31	4	US-09-442-989-3	Sequence 3, Appl	222	26	26.8	46	3	US-09-100-600A-44	Sequence 44, Appl
150	27	27.8	31	4	US-09-442-989-4	Sequence 4, Appl	223	26	26.8	46	3	US-09-100-600A-48	Sequence 48, Appl
151	27	27.8	31	4	US-09-442-989-5	Sequence 5, Appl	224	26	26.8	46	3	US-09-230-548-25	Sequence 25, Appl
152	27	27.8	31	4	US-09-442-989-6	Sequence 6, Appl	225	25.5	26.3	46	3	US-09-040-738-4	Sequence 4, Appl
153	27	27.8	31	4	US-09-442-989-7	Sequence 7, Appl	226	25.5	26.3	39	3	US-08-652-428A-4	Sequence 4, Appl
154	27	27.8	31	4	US-09-442-989-8	Sequence 8, Appl	227	25.5	26.3	39	3	US-07-956-700B-12	Sequence 12, Appl
155	27	27.8	31	4	US-09-442-989-9	Sequence 9, Appl	228	25	25.8	20	1	US-07-956-700B-26	Sequence 26, Appl
156	27	27.8	31	4	US-09-442-989-10	Sequence 10, Appl	229	25	25.8	20	1	US-08-218-025A-71	Sequence 71, Appl
157	27	27.8	31	4	US-09-442-989-11	Sequence 11, Appl	230	25	25.8	20	1	US-08-306-116A-28	Sequence 28, Appl
158	27	27.8	31	4	US-09-442-989-12	Sequence 12, Appl	231	25	25.8	20	1	US-08-476-537-12	Sequence 12, Appl
159	27	27.8	31	4	US-09-442-989-17	Sequence 17, Appl	232	25	25.8	20	1	US-08-476-537-26	Sequence 26, Appl
160	27	27.8	31	4	US-09-442-989-32	Sequence 32, Appl	233	25	25.8	20	1	US-08-485-607-12	Sequence 12, Appl
161	27	27.8	34	4	US-09-228-990-46	Sequence 46, Appl	234	25	25.8	20	1	US-08-485-607-26	Sequence 26, Appl
162	27	27.8	34	4	US-09-442-989-18	Sequence 18, Appl	235	25	25.8	20	2	US-08-238-821B-48	Sequence 48, Appl
163	27	27.8	34	4	US-09-442-989-46	Sequence 46, Appl	236	25	25.8	20	2	US-08-475-879-12	Sequence 12, Appl
164	27	27.8	35	3	US-09-082-279B-714	Sequence 714, App	237	25	25.8	20	2	US-08-475-879-26	Sequence 26, Appl
165	27	27.8	35	3	US-09-315-304B-714	Sequence 714, App	238	25	25.8	20	2	US-08-475-879-12	Sequence 12, Appl
166	27	27.8	35	4	US-09-834-784-714	Sequence 714, App	239	25	25.8	20	4	US-09-834-759-537	Sequence 537, App
167	27	27.8	35	4	US-09-515-965A-714	Sequence 714, App	240	25	25.8	20	4	US-09-834-759-547	Sequence 547, App
168	27	27.8	35	4	US-09-350-641C-714	Sequence 714, App	241	25	25.8	20	5	PCT-US95-0574A-48	Sequence 48, Appl
169	27	27.8	35	4	US-09-350-841A-714	Sequence 714, App	242	25	25.8	20	5	PCT-US95-0574A-48	Sequence 166, App
170	27	27.8	36	4	US-09-350-841A-714	Sequence 123, App	243	25	25.8	21	2	US-08-934-915-166	Sequence 19, Appl
171	27	27.8	37	4	US-09-106-568E-123	Sequence 123, App	244	25	25.8	21	3	US-09-399-494-19	Sequence 19, Appl
172	27	27.8	40	2	US-09-170-767-56669	Sequence 56669, A	245	25	25.8	21	3	US-09-399-494-19	Sequence 19, Appl
173	27	27.8	40	3	US-08-939-002A-2	Sequence 2, Appl	246	25	25.8	21	3	US-09-399-494-19	Sequence 19, Appl

247	25	25.8	21	3	US-09-399-494-21	Sequence 21, Appl	320	25	25.8	36	1	US-08-302-485-12	Sequence 12, Appl
248	25	25.8	22	1	US-07-752-101A-5	Sequence 5, Appl	321	25	25.8	36	1	US-08-302-485-13	Sequence 13, Appl
249	25	25.8	22	2	US-08-934-915-4	Sequence 4, Appl	322	25	25.8	36	2	US-08-487-074-1	Sequence 1, Appl
250	25	25.8	22	4	US-09-466-384-7	Sequence 7, Appl	323	25	25.8	36	2	US-08-487-074-2	Sequence 2, Appl
251	25	25.8	23	1	US-08-290-373B-2	Sequence 2, Appl	324	25	25.8	36	2	US-08-487-074-3	Sequence 3, Appl
252	25	25.8	23	2	US-08-487-074-12	Sequence 12, Appl	325	25	25.8	36	2	US-08-478-435-28	Sequence 28, Appl
253	25	25.8	23	3	US-08-468-011A-27	Sequence 27, Appl	326	25	25.8	36	2	US-08-337-483-28	Sequence 28, Appl
254	25	25.8	23	3	US-08-763-458-14	Sequence 14, Appl	327	25	25.8	36	2	US-08-478-373-28	Sequence 28, Appl
255	25	25.8	23	3	US-09-236-468A-27	Sequence 27, Appl	328	25	25.8	36	3	US-08-474-671-28	Sequence 28, Appl
256	25	25.8	23	3	US-08-986-627-14	Sequence 14, Appl	329	25	25.8	36	3	US-08-483-577A-28	Sequence 28, Appl
257	25	25.8	23	4	US-09-786-569-22	Sequence 22, Appl	330	25	25.8	36	3	US-08-763-458-1	Sequence 1, Appl
258	25	25.8	24	1	US-09-323-854-14	Sequence 14, Appl	331	25	25.8	36	3	US-08-763-458-2	Sequence 2, Appl
259	25	25.8	24	1	US-07-752-101A-1	Sequence 1, Appl	332	25	25.8	36	3	US-08-763-458-3	Sequence 3, Appl
260	25	25.8	25	2	US-08-487-074-5	Sequence 5, Appl	333	25	25.8	36	3	US-08-679-006-14	Sequence 14, Appl
261	25	25.8	25	3	US-08-763-458-5	Sequence 5, Appl	334	25	25.8	36	3	US-09-127-680-12	Sequence 12, Appl
262	25	25.8	25	3	US-09-028-937-35	Sequence 35, Appl	335	25	25.8	36	3	US-08-897-438-28	Sequence 28, Appl
263	25	25.8	25	3	US-09-227-357-464	Sequence 464, App	336	25	25.8	36	3	US-09-044-098A-11	Sequence 11, Appl
264	25	25.8	25	3	US-08-986-627-5	Sequence 5, Appl	337	25	25.8	36	3	US-09-044-098A-12	Sequence 12, Appl
265	25	25.8	25	4	US-09-323-854-5	Sequence 5, Appl	338	25	25.8	36	3	US-09-044-098A-13	Sequence 13, Appl
266	25	25.8	28	3	US-08-679-006-15	Sequence 15, Appl	339	25	25.8	36	3	US-08-986-627-1	Sequence 1, Appl
267	25	25.8	28	3	US-09-040-485-3	Sequence 3, Appl	340	25	25.8	36	3	US-08-986-627-2	Sequence 2, Appl
268	25	25.8	28	4	US-09-635-872A-27	Sequence 27, Appl	341	25	25.8	36	3	US-08-986-627-3	Sequence 3, Appl
269	25	25.8	28	4	US-09-636-077A-27	Sequence 27, Appl	342	25	25.8	36	3	US-08-986-627-24	Sequence 24, Appl
270	25	25.8	28	4	US-09-636-060C-27	Sequence 27, Appl	343	25	25.8	36	3	US-08-637-654-28	Sequence 28, Appl
271	25	25.8	28	4	US-09-659-521-3	Sequence 3, Appl	344	25	25.8	36	3	US-08-637-654-28	Sequence 28, Appl
272	25	25.8	28	4	US-09-398-405-15	Sequence 15, Appl	345	25	25.8	36	4	US-09-728-484-11	Sequence 11, Appl
273	25	25.8	28	4	US-09-636-596C-27	Sequence 27, Appl	346	25	25.8	36	4	US-09-728-484-12	Sequence 12, Appl
274	25	25.8	28	4	US-08-857-389-11	Sequence 11, Appl	347	25	25.8	36	4	US-09-728-484-13	Sequence 13, Appl
275	25	25.8	29	4	US-08-303-270-6	Sequence 6, Appl	348	25	25.8	36	4	US-09-323-854-1	Sequence 1, Appl
276	25	25.8	30	1	US-08-487-074-4	Sequence 4, Appl	349	25	25.8	36	4	US-09-323-854-2	Sequence 2, Appl
277	25	25.8	30	2	US-08-487-074-13	Sequence 13, Appl	350	25	25.8	36	4	US-09-323-854-3	Sequence 3, Appl
278	25	25.8	30	3	US-08-763-458-4	Sequence 4, Appl	351	25	25.8	36	4	US-09-398-405-14	Sequence 14, Appl
279	25	25.8	30	3	US-08-763-458-15	Sequence 15, Appl	352	25	25.8	37	1	US-08-277-540-1	Sequence 1, Appl
280	25	25.8	30	3	US-08-986-627-4	Sequence 4, Appl	353	25	25.8	37	1	US-08-190-802A-225	Sequence 225, App
281	25	25.8	30	3	US-08-986-627-15	Sequence 15, Appl	354	25	25.8	37	1	US-08-430-787A-1	Sequence 1, Appl
282	25	25.8	30	3	US-09-323-854-4	Sequence 4, Appl	355	25	25.8	37	3	US-08-477-346-225	Sequence 225, App
283	25	25.8	30	4	US-09-323-854-15	Sequence 15, Appl	356	25	25.8	37	3	US-08-487-072A-225	Sequence 225, App
284	25	25.8	30	4	US-09-622-782-3	Sequence 3, Appl	357	25	25.8	37	4	US-09-270-767-40635	Sequence 40635, A
285	25	25.8	31	3	US-09-227-357-341	Sequence 341, App	358	25	25.8	37	4	Sequence 42216, A	Sequence 42216, A
286	25	25.8	32	3	US-08-935-263-21	Sequence 21, Appl	359	25	25.8	37	4	Sequence 55851, A	Sequence 55851, A
287	25	25.8	33	3	US-09-082-279B-367	Sequence 367, App	360	25	25.8	42	4	Sequence 55644, A	Sequence 55644, A
288	25	25.8	33	3	US-09-082-279B-368	Sequence 368, App	361	25	25.8	42	4	Sequence 14, Appl	Sequence 14, Appl
289	25	25.8	33	3	US-09-082-279B-369	Sequence 369, App	362	25	25.8	43	1	Sequence 15, Appl	Sequence 15, Appl
290	25	25.8	33	3	US-09-315-304B-367	Sequence 367, App	363	25	25.8	43	1	Sequence 85, Appl	Sequence 85, Appl
291	25	25.8	33	3	US-09-315-304B-368	Sequence 368, App	364	25	25.8	43	1	Sequence 85, Appl	Sequence 85, Appl
292	25	25.8	33	3	US-09-315-304B-369	Sequence 369, App	365	25	25.8	43	1	Sequence 85, Appl	Sequence 85, Appl
293	25	25.8	33	3	US-09-594-185-21	Sequence 21, Appl	366	25	25.8	43	1	Sequence 36, Appl	Sequence 36, Appl
294	25	25.8	33	4	US-09-834-784-367	Sequence 367, App	367	25	25.8	43	1	Sequence 36, Appl	Sequence 36, Appl
295	25	25.8	33	4	US-09-834-784-368	Sequence 368, App	368	25	25.8	43	1	Sequence 36, Appl	Sequence 36, Appl
296	25	25.8	33	4	US-09-834-784-369	Sequence 369, App	369	25	25.8	43	1	Sequence 36, Appl	Sequence 36, Appl
297	25	25.8	33	4	US-09-515-965A-367	Sequence 367, App	370	25	25.8	43	1	Sequence 36, Appl	Sequence 36, Appl
298	25	25.8	33	4	US-09-515-965A-368	Sequence 368, App	371	25	25.8	43	1	Sequence 36, Appl	Sequence 36, Appl
299	25	25.8	33	4	US-09-515-965A-369	Sequence 369, App	372	25	25.8	43	2	Sequence 36, Appl	Sequence 36, Appl
300	25	25.8	33	4	US-09-350-641C-367	Sequence 367, App	373	25	25.8	43	2	Sequence 36, Appl	Sequence 36, Appl
301	25	25.8	33	4	US-09-350-641C-368	Sequence 368, App	374	25	25.8	43	2	Sequence 36, Appl	Sequence 36, Appl
302	25	25.8	33	4	US-09-350-641C-369	Sequence 369, App	375	25	25.8	43	2	Sequence 36, Appl	Sequence 36, Appl
303	25	25.8	33	4	US-10-033-078-21	Sequence 21, Appl	376	25	25.8	43	5	Sequence 36, Appl	Sequence 36, Appl
304	25	25.8	33	4	US-09-350-841A-367	Sequence 367, App	377	25	25.8	43	5	Sequence 36, Appl	Sequence 36, Appl
305	25	25.8	33	4	US-09-350-841A-368	Sequence 368, App	378	25	25.8	43	5	Sequence 36, Appl	Sequence 36, Appl
306	25	25.8	33	4	US-09-350-841A-369	Sequence 369, App	379	25	25.8	43	5	Sequence 36, Appl	Sequence 36, Appl
307	25	25.8	34	1	US-08-014-153D-30	Sequence 30, Appl	380	25	25.8	45	3	Sequence 1, Appl	Sequence 1, Appl
308	25	25.8	34	3	US-09-007-905-19	Sequence 19, Appl	381	25	25.8	45	3	Sequence 1, Appl	Sequence 1, Appl
309	25	25.8	34	3	US-08-142-551B-45	Sequence 45, Appl	382	25	25.8	46	1	Sequence 1, Appl	Sequence 1, Appl
310	25	25.8	34	1	US-08-196-748-2	Sequence 2, Appl	383	25	25.8	47	4	Sequence 4, Appl	Sequence 4, Appl
311	25	25.8	35	2	US-09-493-795B-248	Sequence 248, App	384	25	25.8	47	4	Sequence 40584, A	Sequence 40584, A
312	25	25.8	35	3	US-08-229-009-5	Sequence 5, Appl	385	25	25.8	48	4	Sequence 55800, A	Sequence 55800, A
313	25	25.8	35	3	US-08-487-890A-28	Sequence 28, Appl	386	25	25.8	49	4	Sequence 254, App	Sequence 254, App
314	25	25.8	36	1	US-08-302-485-11	Sequence 11, Appl	387	25	25.8	49	4	Sequence 209, App	Sequence 209, App
315	25	25.8	36	1			388	25	25.8	49	4	Sequence 210, App	Sequence 210, App
316	25	25.8	36	1			389	25	25.8	49	4	Sequence 211, App	Sequence 211, App
317	25	25.8	36	1			390	25	25.8	49	4	Sequence 212, App	Sequence 212, App
318	25	25.8	36	1			391	25	25.8	50	1	Sequence 213, App	Sequence 213, App
319	25	25.8	36	1			392	25	25.8	50	2	Sequence 11, Appl	Sequence 11, Appl
												Sequence 46, Appl	Sequence 46, Appl

393 25 25.8 50 2 US-08-757-536-46 Sequence 46, Appl
394 25 25.8 50 3 US-09-314-093-46 Sequence 46, Appl
395 25 25.8 50 3 US-09-250-848-46 Sequence 46, Appl
396 25 25.8 50 3 US-09-251-885-46 Sequence 46, Appl
397 25 25.8 50 3 US-09-337-635-46 Sequence 46, Appl
398 25 25.8 50 4 US-09-337-280-46 Sequence 251, App
399 25 25.8 50 4 US-09-601-802D-251 Sequence 252, App
400 25 25.8 50 4 US-09-601-802D-252 Sequence 253, App
401 25 25.8 50 4 US-09-601-802D-253 Sequence 255, App
402 25 25.8 50 4 US-09-601-802D-255 Sequence 9, Appli
403 24.5 25.3 23 3 US-09-029-348-9 Sequence 3, Appli
404 24.5 25.3 23 4 US-09-585-887-3 Sequence 61, Appl
405 24.5 25.3 44 3 US-09-283-578-3 Sequence 28, Appl
406 24.5 25.3 20 1 US-08-488-212A-28 Sequence 28, Appl
407 24 24.7 20 2 US-08-320-306-28 Sequence 28, Appl
408 24 24.7 20 2 US-08-488-209B-28 Sequence 4, Appli
409 24 24.7 20 2 US-08-408-011-28 Sequence 12, Appl
410 24 24.7 20 3 US-09-308-935-4 Sequence 39, Appl
411 24 24.7 20 4 US-09-142-54D-12 Sequence 9, Appli
412 24 24.7 21 2 US-08-493-235-39 Sequence 132, App
413 24 24.7 22 4 US-10-053-485-9 Sequence 15, Appl
414 24 24.7 23 3 US-08-485-324-15 Sequence 15, Appl
415 24 24.7 23 3 US-08-447-506-15 Sequence 15, Appl
416 24 24.7 23 3 US-08-235-437-15 Sequence 15, Appl
417 24 24.7 23 3 US-08-407-252-4 Sequence 6, Appli
418 24 24.7 24 4 US-09-206-576-6 Sequence 240, App
419 24 24.7 24 4 US-09-721-108-240 Sequence 19, Appl
420 24 24.7 25 1 US-08-172-331B-8 Sequence 8, Appli
421 24 24.7 25 1 US-08-258-852-7 Sequence 7, Appli
422 24 24.7 25 2 US-09-170-754B-7 Sequence 25, Appl
423 24 24.7 25 3 US-08-455-625-25 Sequence 25, Appl
424 24 24.7 25 3 US-08-455-685-25 Sequence 25, Appl
425 24 24.7 25 3 US-08-060-988A-25 Sequence 7, Appli
426 24 24.7 25 4 US-09-450-071A-7 Sequence 39, Appl
427 24 24.7 25 4 US-09-386-959-39 Sequence 43251, A
428 24 24.7 25 4 US-09-270-767-43251 Sequence 58361, A
429 24 24.7 25 5 PCT-US94-05142-25 Sequence 25, Appl
430 24 24.7 25 5 PCT-US95-07539-7 Sequence 7, Appli
431 24 24.7 26 3 US-08-335-733D-1 Sequence 1, Appli
432 24 24.7 28 3 US-09-099-041A-22 Sequence 22, Appl
433 24 24.7 28 3 US-09-245-281-22 Sequence 22, Appl
434 24 24.7 28 4 US-09-207-359B-22 Sequence 22, Appl
435 24 24.7 28 4 US-09-340-620A-22 Sequence 22, Appl
436 24 24.7 28 4 US-09-865-364-22 Sequence 3, Appli
437 24 24.7 29 1 US-08-328-152A-3 Sequence 29, Appl
438 24 24.7 29 2 US-08-588-258B-29 Sequence 29, Appl
439 24 24.7 29 3 US-08-460-505-29 Sequence 139, App
440 24 24.7 29 5 PCT-US96-08295-29 Sequence 39, Appl
441 24 24.7 30 1 US-08-488-252-37 Sequence 37, Appl
442 24 24.7 30 2 US-08-470-720-9 Sequence 9, Appli
443 24 24.7 30 4 US-09-270-767-58742 Sequence 58742, A
444 24 24.7 30 4 US-08-070-455-9 Sequence 9, Appli
445 24 24.7 31 4 US-09-228-990-19 Sequence 19, Appl
446 24 24.7 31 4 US-09-205-258-459 Sequence 459, App
447 24 24.7 31 4 US-09-270-767-37936 Sequence 37936, A
448 24 24.7 31 4 US-09-270-767-53153 Sequence 53153, A
449 24 24.7 31 4 US-09-142-524D-5 Sequence 5, Appli
450 24 24.7 32 4 US-08-737-109-23 Sequence 23, Appl
451 24 24.7 32 6 5204326-123 Patent No. 5204326
452 24 24.7 33 1 US-08-257-528B-83 Sequence 83, Appl
453 24 24.7 33 1 US-08-257-528B-84 Sequence 84, Appl
454 24 24.7 33 1 US-08-460-602A-83 Sequence 83, Appl
455 24 24.7 33 1 US-08-460-602A-84 Sequence 84, Appl
456 24 24.7 33 1 US-08-463-966A-83 Sequence 83, Appl
457 24 24.7 33 1 US-08-463-966A-84 Sequence 83, Appl
458 24 24.7 33 1 US-08-465-217A-83 Sequence 83, Appl
459 24 24.7 33 1 US-08-465-217A-84 Sequence 84, Appl
460 24 24.7 33 1 US-08-465-217A-84 Sequence 84, Appl
461 24 24.7 33 1 US-08-465-217A-84 Sequence 84, Appl
462 24 24.7 33 1 US-08-465-217A-84 Sequence 84, Appl
463 24 24.7 33 1 US-08-465-217A-84 Sequence 84, Appl
464 24 24.7 33 1 US-08-465-217A-84 Sequence 84, Appl
465 24 24.7 33 1 US-08-465-217A-84 Sequence 84, Appl

466 24 24.7 33 2 US-08-464-329A-83 Sequence 83, Appl
467 24 24.7 33 2 US-08-464-329A-84 Sequence 84, Appl
468 24 24.7 33 2 US-08-461-990B-5 Sequence 5, Appli
469 24 24.7 33 2 US-08-462-507A-83 Sequence 83, Appl
470 24 24.7 33 2 US-08-462-507A-84 Sequence 84, Appl
471 24 24.7 33 2 US-08-467-881A-83 Sequence 83, Appl
472 24 24.7 33 2 US-08-467-881A-84 Sequence 84, Appl
473 24 24.7 33 2 US-08-508-930D-44 Sequence 44, Appl
474 24 24.7 33 4 US-09-536-977-18 Sequence 18, Appl
475 24 24.7 33 4 US-08-985-056-66 Sequence 66, Appl
476 24 24.7 34 4 US-09-270-767-58505 Sequence 58505, A
477 24 24.7 34 4 US-09-270-767-58985 Sequence 58985, A
478 24 24.7 35 2 US-08-142-551B-67 Sequence 67, Appl
479 24 24.7 35 3 US-09-082-279B-716 Sequence 716, App
480 24 24.7 35 3 US-09-315-304B-716 Sequence 716, App
481 24 24.7 35 4 US-03-834-784-716 Sequence 716, App
482 24 24.7 35 4 US-09-515-965A-716 Sequence 716, App
483 24 24.7 35 4 US-09-350-641C-716 Sequence 716, App
484 24 24.7 35 4 US-09-270-767-40552 Sequence 40552, A
485 24 24.7 35 4 US-09-270-767-55768 Sequence 55768, A
486 24 24.7 35 4 US-09-350-841A-716 Sequence 716, App
487 24 24.7 36 4 US-09-257-179-54 Sequence 54, Appl
488 24 24.7 36 4 US-09-462-917A-90 Sequence 90, Appl
489 24 24.7 36 4 US-09-790-497A-18 Sequence 18, Appl
490 24 24.7 37 3 US-08-751-344B-16 Sequence 16, Appl
491 24 24.7 37 3 US-08-281-313-5 Sequence 5, Appli
492 24 24.7 37 3 US-08-146-028-18 Sequence 18, Appl
493 24 24.7 38 2 US-08-928-692-19 Sequence 19, Appl
494 24 24.7 38 3 US-08-723-425A-18 Sequence 18, Appl
495 24 24.7 38 3 US-09-112-206-18 Sequence 18, Appl
496 24 24.7 38 3 US-09-339-972-19 Sequence 19, Appl
497 24 24.7 38 4 US-09-576-824A-18 Sequence 18, Appl
498 24 24.7 38 4 US-09-680-497-18 Sequence 18, Appl
499 24 24.7 39 2 US-09-455-625-3 Sequence 3, Appli
500 24 24.7 39 2 US-08-455-625-29 Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-08-475-955-55
; Sequence 55, Application US/08475955
; Patent No. 6641813
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
; TITLE OF INVENTION: AUTOANTIBODIES
; NUMBER OF SEQUENCES: 218
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,955
; FILING DATE: June 7, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/867,819
; FILING DATE: April 13, 1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/648,205
; FILING DATE: January 31, 1991
; PRIOR APPLICATION DATA:


```

; APPLICATION NUMBER: 07/472,947
; FILING DATE: January 31, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMR114CIP(2)DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 1..14
; US-08-475-955-55

Query Match 39.2%; Score 38; DB 4; Length 22;
Best Local Similarity 53.3%; Pred. No. 5;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NHLNSKIAFKIVSQE 17
Db 5 NHLKSKEVKALLQE 19

RESULT 2
US-09-270-767-56675
; Sequence 56675, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56675
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-56675

Query Match 34.0%; Score 33; DB 4; Length 45;
Best Local Similarity 53.3%; Pred. No. 89;
Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 4 HLNSKIAFKIVSQEP 18
Db 25 HYSSEIAFK-KQNP 37

RESULT 3
US-09-270-767-38499
; Sequence 38499, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38499
; LENGTH: 49

; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-38499

Query Match 34.0%; Score 33; DB 4; Length 49;
Best Local Similarity 66.7%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 HLNSKIAFK 12
Db 24 HLNSPVAIK 32

RESULT 4
US-09-270-767-53716
; Sequence 53716, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53716
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-53716

Query Match 34.0%; Score 33; DB 4; Length 49;
Best Local Similarity 66.7%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 HLNSKIAFK 12
Db 24 HLNSPVAIK 32

RESULT 5
US-08-427-072-10
; Sequence 10, Application US/08427072
; Patent No. 5776673
; GENERAL INFORMATION:
; APPLICANT: Tabor, Stanley
; APPLICANT: Richardson, Charles
; TITLE OF INVENTION: TREATMENT AND DETECTION OF
; TITLE OF INVENTION: TUBERCULOSIS, LEPROSY, AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,072
; FILING DATE: April 21, 1995
; CLASSIFICATION: 435

```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 211/294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-427-072-10

Query Match 33.0%; Score 32; DB 1; Length 26;
Best Local Similarity 57.1%; Pred. No. 70;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQEPA 19
Db 6 SKKIQFGIVQESA 19

RESULT 6
US-09-172-841-29
; Sequence 29, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT APPLICATION NUMBER: US/09/172,841
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-172-841-29

Query Match 32.0%; Score 31; DB 3; Length 33;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQ 16
Db 1 LPABITFKIFSQ 12

RESULT 7
US-08-951-621-29
; Sequence 29, Application US/08951621
; Patent No. 6573094
; GENERAL INFORMATION:
; APPLICANT: HARPER, JEFFREY W.
; APPLICANT: ELLEDGE, STEPHEN J.
; TITLE OF INVENTION: F-BOX GENES AND PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: BCM-02999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-951-621-29

Query Match 32.0%; Score 31; DB 4; Length 33;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQ 16
Db 1 LPABITFKIFSQ 12

RESULT 8
US-09-690-454-72
; Sequence 72, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
```

```

; LOCATION: (34)
; OTHER INFORMATION: Xaa equals stop translation
US-09-690-454-72

Query Match      32.0%; Score 31; DB 4; Length 34;
Best Local Similarity 47.4%; Pred. No. 1.4e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY      5 LNSKIAFKIV---SQEPA 19
      |||||: :|||: |||
Db      11 LNSKLVAAVNKLKASQMPA 29

RESULT 9
US-09-270-767-36721
; Sequence 36721, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36721
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-36721

Query Match      32.0%; Score 31; DB 4; Length 46;
Best Local Similarity 29.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 EPNHLSKIAPKIVSQE 17
      :|||: |||: :
Db      19 DPRYLMENITYKIFNNQ 35

RESULT 10
US-09-270-767-51938
; Sequence 51938, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51938
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-51938

Query Match      32.0%; Score 31; DB 4; Length 46;
Best Local Similarity 29.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 EPNHLSKIAPKIVSQE 17
      :|||: |||: :
Db      19 DPRYLMENITYKIFNNQ 35

RESULT 11
US-09-288-143-87
; Sequence 87, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.

```

```

; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P1
; CURRENT APPLICATION NUMBER: US/09/288,143
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/US98/21142
; EARLIER FILING DATE: 1998-10-08
; EARLIER APPLICATION NUMBER: 60/061,463
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,529
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/071,498
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,527
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,536
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,532
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals stop translation
US-09-288-143-87

Query Match      32.0%; Score 31; DB 4; Length 48;
Best Local Similarity 45.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      9 IAEKIVSQEPA 19
      :|||: :|||
Db      19 VAEFLINQIPA 29

RESULT 12
US-09-790-497A-445
; Sequence 445, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 445
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-445

Query Match      30.9%; Score 30; DB 4; Length 20;

```

```
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 9 IAFKIVSQE 17
   :||||:|
Db 4 VAFKIMSGE 12

RESULT 13
US-09-576-824A-445
; Sequence 445, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 445
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-576-824A-445

Query Match 30.9%; Score 30; DB 4; Length 20;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 9 IAFKIVSQE 17
   :||||:|
Db 4 VAFKIMSGE 12

RESULT 14
US-09-205-258-898
; Sequence 898, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 898
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-898
```


REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: DC-0050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (856) 810-1515
TELEFAX: (856) 810-1454
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 33
TYPE: Amino acid
TOPOLOGY: Linear
US-09-101-146-15

Query Match 30.9%; Score 30; DB 3; Length 33;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLNSKIA 10
|:|:|
Db 12 HNNAKIA 18

RESULT 17
US-09-194-296-13
; Sequence 13, Application US/09194296
; Patent No. 6486125
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA
; TITLE OF INVENTION: SYNTHETIC APPROACH TO DESIGNED CHEMICAL
; STRUCTURES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
; STREET: 119 No. 6486125th Fourth Street, Suite 203
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: U.S.A.
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/194,296
; FILING DATE: 15-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/653,632
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCORMACK, MYRA M.
; REGISTRATION NUMBER: 36,602
; REFERENCE/DOCKET NUMBER: 110.00330220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1225
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-194-296-13

Query Match 30.9%; Score 30; DB 4; Length 33;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HLNSKIAFKI 13
|:|:|:|
Db 15 HLKWKIIFKL 24

RESULT 18
US-09-149-476-545
; Sequence 545, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580

; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 30.9%; Score 30; DB 4; Length 39;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIV 14
|||:|
Db 27 NHLAFRILFFIV 38

RESULT 19
US-09-149-476-381
; Sequence 381, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585

EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 30.9%; Score 30; DB 4; Length 40;
Best Local Similarity 58.3%; Pred. No. 2.6e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 NHLNFKIAFKIV 14
Db 27 NHLAFILFIV 38

RESULT 20
US-08-821-619-10
Sequence 10, Application US/08821619
Patent No. 5945326
GENERAL INFORMATION:
APPLICANT: Morgan, Richard
APPLICANT: Chang, Zhiyuh
APPLICANT: Mersha, Fana
TITLE OF INVENTION: METHOD FOR CLONING AND
PRODUCING THE SpeI RESTRICTION ENDONUCLEASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: New England Biolabs, Inc.
STREET: 32 Tozer Road
CITY: Beverly
STATE: MA
COUNTRY: US
ZIP: 01915
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/821,619
FILING DATE: 20-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Gregory D
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-927-5054
TELEFAX: 509-927-1705
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-821-619-10

Query Match 29.9%; Score 29; DB 2; Length 28;
Best Local Similarity 55.6%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKI 9
Db 4 DPNKLSAL 12

RESULT 21
US-07-829-462-3
Sequence 3, Application US/07829462
Patent No. 5453489
GENERAL INFORMATION:
APPLICANT: Ruoslahti, Erkki I.
APPLICANT: Morla, Alex
TITLE OF INVENTION: NOVEL FRAGMENTS OF FIBRONECTIN AND
METHODS TO MODULATE THE ROLE OF FIBRONECTIN IN
EXTRACELLULAR MATRIX ASSEMBLY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/829,462
FILING DATE: 19920131
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, CATHRYN A.

```

,      REGISTRATION NUMBER: 31,547
,
,      REFERENCE/DOCKET NUMBER: P-LA 9179
,
,      TELECOMMUNICATION INFORMATION:
,
,      TELEPHONE: 619-535-9001
,      TELEFAX: 619-535-8949
,      INFORMATION FOR SEQ ID NO: 3:
,      SEQUENCE CHARACTERISTICS:
,      LENGTH: 31 amino acids
,      TYPE: AMINO ACID
,      TOPOLOGY: linear
,
,      US-07-829,462-3

```

Query Match 29.9%; Score 29; DB 1; Length 31;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5: Conservative 0; Mismatches 1; Indels

```

RESULT 22
US-08-340-812-3
; Sequence 3, Application US/08340812
; Patent No. 5629291
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND METHODS OF
; MODULATING FIBRONECTIN EXTRACELLULAR MATRIX ASSEMBLY
; TITLE OF INVENTION: MODULATING FIBRONECTIN EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340-812

```

Query Match 29.9%; Score 29; DB 1; Length 31;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels

RESULT 23
US-08-459-064B-3
; Sequence 3, Application US/08459064B
; Patent No. 5747452
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: A METHOD OF MODULATING TUMOR CELL MIGRATION
; USING FIBRONECTIN TYPE III PEPTIDES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,064B
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,462
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,626
; FILING DATE: 16-FEB-1993
; PRIOR APPLICATION DATA:

```
Query Match      29.9%; Score 29; DB 1; Length 31;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5: Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

RESULT 24
US-08-460-421A-3
; Sequence 3, Application US/08460421A
; Patent No. 5837813
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND METHODS OF
; TITLE OF INVENTION: MODULATING FIBRONECTIN EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSES: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700

;/ CITY: SAN DIEGO
;/ STATE: CALIFORNIA
;/ COUNTRY: UNITED STATES
;/ ZIP: 92122
;/ COMPUTER READABLE FORM:
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/460,421A
;/ FILING DATE: 01-JUN-1995
;/ CLASSIFICATION: 530
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 07/829,462
;/ FILING DATE: 31-JAN-1992
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 08/021,626
;/ FILING DATE: 16-FEB-1993
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 08/340,812
;/ FILING DATE: 17-NOV-1994
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: CAMPBELL, CATHRYN A.
;/ REGISTRATION NUMBER: 31,815
;/ REFERENCE/DOCKET NUMBER: P-LA 1542
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 619-535-9001
;/ TELEFAX: 619-535-8949
;/ INFORMATION FOR SEQ ID NO: 3:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 31 amino acids
;/ TYPE: amino acid
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: peptide
;/ US-08-460-421A-3

Query Match 29.9%; Score 29; DB 2; Length 31;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLS 7
Db 5 PGLHS 10

RESULT 25
PCT-US93-00909-3
;/ SEQUENCE 3, Application PC/TUS9300909
;/ GENERAL INFORMATION:
;/ APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION
;/ TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND METHODS
;/ TITLE OF INVENTION: OF MODULATING FIBRONECTIN
;/ TITLE OF INVENTION: EXTRACELLULAR MATRIX ASSEMBLY
;/ NUMBER OF SEQUENCES: 16
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: CAMPBELL AND FLORES
;/ STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
;/ CITY: SAN DIEGO
;/ STATE: CALIFORNIA
;/ COUNTRY: UNITED STATES
;/ ZIP: 92122
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: PCT/US93/00909
;/ FILING DATE: 19930201
;/ CLASSIFICATION:
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: BELLAS, CHRISTINE M.

;/ REGISTRATION NUMBER: 34,122
;/ REFERENCE/DOCKET NUMBER: FP-LA 9462
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 619-535-9001
;/ TELEFAX: 619-535-8949
;/ INFORMATION FOR SEQ ID NO: 3:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 31 amino acids
;/ TYPE: AMINO ACID
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: peptide
;/ PCT-US93-00909-3

Query Match 29.9%; Score 29; DB 5; Length 31;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLS 7
Db 5 PGLHS 10

RESULT 26
US-09-036-315-22
;/ SEQUENCE 22, Application US/09036315
;/ Patent No. 6218523
;/ GENERAL INFORMATION:
;/ APPLICANT: French, Cynthia K.
;/ APPLICANT: Schneider, Patrick A.
;/ APPLICANT: Yamamoto, Karen K.
;/ TITLE OF INVENTION: Prostate Cancer-Specific Marker
;/ NUMBER OF SEQUENCES: 27
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Townsend and Townsend and Crew LLP
;/ STREET: Two Embarcadero Center, Eighth Floor
;/ CITY: San Francisco
;/ STATE: California
;/ COUNTRY: USA
;/ ZIP: 94111-3834
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ FILING DATE: 06-MAR-1998
;/ CLASSIFICATION: 435
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 60/047,811
;/ FILING DATE: 15-MAY-1997
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 60/041,246
;/ FILING DATE: 07-MAR-1997
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Storella, John R.
;/ REGISTRATION NUMBER: 32,944
;/ REFERENCE/DOCKET NUMBER: 018002-000210US
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (415) 576-0200
;/ TELEFAX: (415) 576-0300
;/ INFORMATION FOR SEQ ID NO: 22:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 34 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS:
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: peptide
;/ US-09-036-315-22

Query Match 29.9%; Score 29; DB 3; Length 34;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
;
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 16
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-515-965A-1908

Query Match      29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NHLNSKIAFKIVSQE 17
Db 14 NKXNGTDAVKLIKQE 28

RESULT 29
US-09-515-965A-1909
; Sequence 1909, Application US/09515965A
; Patent No. 6623741
; GENERAL INFORMATION:
; APPLICANT: Antczak, J.
; APPLICANT: Delmedico, M.
; APPLICANT: Erickson, J.
; APPLICANT: Lambert, D.
; APPLICANT: Sista, P.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE FUSION-
; FILE REFERENCE: 7872-073
; CURRENT APPLICATION NUMBER: US/09/515,965A
; CURRENT FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1994
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1909
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 15
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-515-965A-1909

Query Match      29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NHLNSKIAFKIVSQE 17
Db 13 NKXNGTDAVKLIKQE 27

RESULT 30
US-09-350-641C-1670
; Sequence 1670, Application US/09350641C
; Patent No. 6656906
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-067
; CURRENT APPLICATION NUMBER: US/09/350,641C
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
```

```
Qy 2 PNHLNSKIAF 11
Db 23 PENLNSKKKF 32

RESULT 27
US-09-515-965A-1907
; Sequence 1907, Application US/09515965A
; Patent No. 6623741
; GENERAL INFORMATION:
; APPLICANT: Antczak, J.
; APPLICANT: Delmedico, M.
; APPLICANT: Erickson, J.
; APPLICANT: Lambert, D.
; APPLICANT: Sista, P.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE FUSION-
; FILE REFERENCE: 7872-073
; CURRENT APPLICATION NUMBER: US/09/515,965A
; CURRENT FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1994
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1907
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 17
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-515-965A-1907

Query Match      29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NHLNSKIAFKIVSQE 17
Db 15 NKXNGTDAVKLIKQE 29

RESULT 28
US-09-515-965A-1908
; Sequence 1908, Application US/09515965A
; Patent No. 6623741
; GENERAL INFORMATION:
; APPLICANT: Antczak, J.
; APPLICANT: Delmedico, M.
; APPLICANT: Erickson, J.
; APPLICANT: Lambert, D.
; APPLICANT: Sista, P.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE FUSION-
; FILE REFERENCE: 7872-073
; CURRENT APPLICATION NUMBER: US/09/515,965A
; CURRENT FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1994
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1908
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```

; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1670
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 17
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-350-641C-1670

```

```

Query Match      29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 3 NHLNSKIAFKIVSQE 17
   | | | | | | | | | |
Db 15 NXKNGTDAVKLIKQE 29

```

RESULT 31

```

US-09-350-641C-1671
; Sequence 1671, Application US/09350641C
; Patent No. 6656906
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Answer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-067

```

```

; CURRENT APPLICATION NUMBER: US/09/350,641C
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1671
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 16
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-350-641C-1671

```

```

Query Match      29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 3 NHLNSKIAFKIVSQE 17
   | | | | | | | | | |
Db 14 NXKNGTDAVKLIKQE 28

```

RESULT 32

```

US-09-350-641C-1672
; Sequence 1672, Application US/09350641C
; Patent No. 6656906
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.

```

```

; APPLICANT: Answer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-067
; CURRENT APPLICATION NUMBER: US/09/350,641C
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1672
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 15
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-350-641C-1672

```

```

Query Match      29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 3 NHLNSKIAFKIVSQE 17
   | | | | | | | | | |
Db 13 NXKNGTDAVKLIKQE 27

```

RESULT 33

```

US-09-350-841A-1859
; Sequence 1859, Application US/09350841A
; Patent No. 6750008
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350,841A
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1859
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 17
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-350-841A-1859

```

```

Query Match      29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 3 NHLNSKIAFKIVSQE 17
   | | | | | | | | | |
Db 15 NXKNGTDAVKLIKQE 29

```

RESULT 34

```

US-09-350-841A-1860
; Sequence 1860, Application US/09350841A
; Patent No. 6750008
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;

```

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350,841A
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1860
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 16
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-350-841A-1860

Query Match 29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQE 17
DB 14 NKXNGTDAVKLIKQE 28

RESULT 35
US-09-350-841A-1861
; Sequence 1861, Application US/09350841A
; Patent No. 6750008
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350,841A
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1861
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
; NAME/KEY: SITE
; LOCATION: 15
; OTHER INFORMATION: Xaa = U (Aminobutyric Acid)
US-09-350-841A-1861

Query Match 29.9%; Score 29; DB 4; Length 34;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQE 17
DB 13 NKXNGTDAVKLIKQE 27

RESULT 36
US-09-177-249-175
; Sequence 175, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Ohad, Nir
; APPLICANT: Kiyosue, Tomohiro
; APPLICANT: Yadegari, Ramin
; APPLICANT: Margossian, Linda
; APPLICANT: Harada, John

; APPLICANT: Goldberg, Robert B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
; TITLE OF INVENTION: Development in Plants
; FILE REFERENCE: 023070-0861200S
; CURRENT APPLICATION NUMBER: US/09/177,249
; CURRENT FILING DATE: 1998-10-22
; EARLIER APPLICATION NUMBER: US 09/071,838
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
; US-09-177-249-175

Query Match 29.9%; Score 29; DB 3; Length 39;
Best Local Similarity 33.3%; Pred. No. 3.8e+02;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIV 14
DB 19 NEVNIRISLIVI 30

RESULT 37
US-08-675-499A-12
; Sequence 12, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-675-499A-12

Query Match 29.9%; Score 29; DB 4; Length 41;
 Best Local Similarity 54.5%; Pred. No. 4.1e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQ 16
 Db 23 DSTIAFDVWSK 33

RESULT 38
 US-08-812-008-12
 ; Sequence 12, Application US/08812008
 ; Patent No. 6602693
 ; GENERAL INFORMATION:
 ; APPLICANT: McDonald, J. A.
 ; APPLICANT: Spicer, A. P.
 ; APPLICANT: Augustine, M. L.
 ; TITLE OF INVENTION: GENE ENCODING HYALURONAN
 ; TITLE OF INVENTION: SYNTHASE
 ; NUMBER OF SEQUENCES: 54
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
 ; STREET: P.O. Box 2938
 ; CITY: Minneapolis
 ; STATE: MN
 ; COUNTRY: U.S.A.
 ; ZIP: 55402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/812,008
 ; FILING DATE: 05-MAR-1997
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 08/675,499
 ; FILING DATE: 03-JUL-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Embretson, Janet E
 ; REGISTRATION NUMBER: 39,665
 ; REFERENCE/DOCKET NUMBER: 150.183US1
 ; TELEPHONE: 612-373-6900
 ; TELEFAX: 612-339-3061
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 41 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-812-008-12

Query Match 29.9%; Score 29; DB 4; Length 41;
 Best Local Similarity 54.5%; Pred. No. 4.1e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 NSKIAFKIVSQ 16
 Db 23 DSTIAFDVWSK 33

RESULT 39
 US-08-897-843A-7
 ; Sequence 7, Application US/08897843A
 ; Patent No. 6514493
 ; GENERAL INFORMATION:
 ; APPLICANT: DeLeo, Albert B.; Loftus, Douglas; Appella, Ettore
 ; TITLE OF INVENTION: CDNA CLONE FOR MURINE TUMOR
 ; TITLE OF INVENTION: REJECTION ANTIGEN GP110 AND TUMOR PEPTIDE VACCINE

; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Diane R. Meyers
 ; STREET: 600 Grant Street, 42nd Floor
 ; CITY: Pittsburgh
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 15219
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/997,843A
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 44
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE:
 ; US-08-897-843A-7

Query Match 29.9%; Score 29; DB 4; Length 44;
 Best Local Similarity 38.9%; Pred. No. 4.4e+02;
 Matches 7; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 2 PNHLSKIAFKIVSQEPA 19
 Db 25 PSHLN--LVFLFLSRAAA 40

RESULT 40
 US-09-205-258-340
 ; Sequence 340, Application US/09205258
 ; Patent No. 6525174
 ; GENERAL INFORMATION:
 ; APPLICANT: Young et al.
 ; TITLE OF INVENTION: 207 Human Secreted Proteins
 ; FILE REFERENCE: P2007P1
 ; CURRENT APPLICATION NUMBER: US/09/205,258
 ; CURRENT FILING DATE: 1998-12-04
 ; EARLIER APPLICATION NUMBER: PCT/US98/11422
 ; EARLIER FILING DATE: 1998-06-04
 ; EARLIER APPLICATION NUMBER: 60/048,885
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,375
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,881
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,880
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,896
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,020
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,876
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,895
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,894
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,894
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,971
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,964
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,882
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,899

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 35038
LENGTH: 49
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35038

Query Match 29.4%; Score 28.5; DB 4; Length 49;
Best Local Similarity 63.6%; Pred. No. 6.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PNHLNSKIAFK 12
||| : |||
Db 30 PNH-HSSIALK 39

RESULT 42
US-09-270-767-50255
Sequence 50255, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 50255
LENGTH: 49
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-50255

Query Match 29.4%; Score 28.5; DB 4; Length 49;
Best Local Similarity 63.6%; Pred. No. 6.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PNHLNSKIAFK 12
||| : |||
Db 30 PNH-HSSIALK 39

RESULT 43
US-08-851-567B-70
Sequence 70, Application US/08851567B
Patent No. 6528484
GENERAL INFORMATION:
APPLICANT: Ensign, Gerald C
APPLICANT: Bowen, David J
APPLICANT: Petell, James
APPLICANT: Fatig, Raymond
APPLICANT: Schoonover, Sue
APPLICANT: french-Constant, Richard
APPLICANT: Rocheleau, Thomas A.
APPLICANT: Blackburn, Michael B.
APPLICANT: Hey, Timothy D.
APPLICANT: Merlo, Donald J.
APPLICANT: Orr, Gregory L.
APPLICANT: Roberts, Jean L.
APPLICANT: Strickland, James A.
APPLICANT: Guo, Lining
APPLICANT: Ciche, Todd A.

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 340
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
US-09-205-258-340

Query Match 29.9%; Score 29; DB 4; Length 46;
Best Local Similarity 37.5%; Pred. No. 4.6e+02;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 NHLNSKIAFKIVSQEP 18
| : : : :
Db 26 NLFTSQIKYIVSEK 41

RESULT 41
US-09-270-767-35038
Sequence 35038, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:

APPLICANT: Sukhapinda, Kitisri
TITLE OF INVENTION: Insecticidal Protein Toxins From Photorhabdus
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: IN
COUNTRY: US
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/851,567B
FILING DATE: 05-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/063,615
FILING DATE: 18-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/395,497
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,255
FILING DATE: 06-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,423
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/705,484
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.93804
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-851-567B-70

Query Match 28.9%; Score 28; DB 4; Length 22;
Best Local Similarity 40.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 PNHLNSKIAPKIVSQ 16
||: :||: ||
Db 3 PNNSSNKLMEYPVYQ 17

RESULT 44
US-09-367-065A-12
Sequence 12, Application US/09367065A
Patent No. 6390821
GENERAL INFORMATION:
APPLICANT: Shokat, Kevan M.
TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
FILE OF INVENTION: Nucleotide Triphosphate Substrates
FILE REFERENCE: 51538-5002-US
CURRENT APPLICATION NUMBER: US/09/367,065A
CURRENT FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: US 08/797,522
PRIOR FILING DATE: 1997-02-07
PRIOR APPLICATION NUMBER: US 60/046,727

PRIOR FILING DATE: 1997-05-16
PRIOR APPLICATION NUMBER: PCT/US98/02522
PRIOR FILING DATE: 1998-02-09
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
OTHER INFORMATION: ATP-binding domain, V323A, I338A
US-09-367-065A-12

Query Match 28.9%; Score 28; DB 3; Length 26;
Best Local Similarity 40.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

QY 6 NSKIA--FKIVSQEP 18
: ||: : ||: ||
Db 2 HEKLAQLYAVVSEEP 16

RESULT 45
US-09-367-065A-13
Sequence 13, Application US/09367065A
Patent No. 6390821
GENERAL INFORMATION:
APPLICANT: Shokat, Kevan M.
TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
FILE OF INVENTION: Nucleotide Triphosphate Substrates
FILE REFERENCE: 51538-5002-US
CURRENT APPLICATION NUMBER: US/09/367,065A
CURRENT FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: US 08/797,522
PRIOR FILING DATE: 1997-02-07
PRIOR APPLICATION NUMBER: US 60/046,727
PRIOR FILING DATE: 1997-05-16
PRIOR APPLICATION NUMBER: PCT/US98/02522
PRIOR FILING DATE: 1998-02-09
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
OTHER INFORMATION: ATP-binding domain, V323A
US-09-367-065A-13

Query Match 28.9%; Score 28; DB 3; Length 26;
Best Local Similarity 40.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

QY 6 NSKIA--FKIVSQEP 18
: ||: : ||: ||
Db 2 HEKLAQLYAVVSEEP 16

RESULT 46
US-09-568-466-12
Sequence 12, Application US/09568466
Patent No. 6521417
GENERAL INFORMATION:
APPLICANT: Shokat, Kevan M.
TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
FILE OF INVENTION: Nucleotide Triphosphate Substrates
FILE REFERENCE: 51538-5002-US
CURRENT APPLICATION NUMBER: US/09/568,466
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: US/09/367,065
PRIOR FILING DATE: 1999-11-17

;; PRIOR APPLICATION NUMBER: US 08/797,522
;; PRIOR FILING DATE: 1997-02-07
;; PRIOR APPLICATION NUMBER: US 60/046,727
;; PRIOR FILING DATE: 1997-05-16
;; PRIOR APPLICATION NUMBER: PCT/US98/02522
;; PRIOR FILING DATE: 1998-02-09
;; NUMBER OF SEQ ID NOS: 20
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 12
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
;; OTHER INFORMATION: ATP-binding domain, V323A, I338A
US-09-568-466-12

Query Match 28.9%; Score 28; DB 4; Length 26;
Best Local Similarity 40.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

QY 6 NSKJA--FKIVSQEP 18
:|:|:|:|:|:|
Db 2 HEKLAQLYAVVSEEP 16

RESULT 47
US-09-568-466-13
;; Sequence 13, Application US/09568466
;; Patent No. 6521417
;; GENERAL INFORMATION:
;; APPLICANT: Shokat, Kevan M.
;; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
;; FILE OF INVENTION: Nucleotide Triphosphate Substrates
;; FILE REFERENCE: 51538-5002-US
;; CURRENT APPLICATION NUMBER: US/09/568,466
;; CURRENT FILING DATE: 2000-05-10
;; PRIOR APPLICATION NUMBER: US/09/367,065
;; PRIOR FILING DATE: 1999-11-17
;; PRIOR APPLICATION NUMBER: US 08/797,522
;; PRIOR FILING DATE: 1997-02-07
;; PRIOR APPLICATION NUMBER: US 60/046,727
;; PRIOR FILING DATE: 1997-05-16
;; PRIOR APPLICATION NUMBER: PCT/US98/02522
;; PRIOR FILING DATE: 1998-02-09
;; NUMBER OF SEQ ID NOS: 20
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 13
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
;; OTHER INFORMATION: ATP-binding domain, V323A
US-09-568-466-13

Query Match 28.9%; Score 28; DB 4; Length 26;
Best Local Similarity 40.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

QY 6 NSKJA--FKIVSQEP 18
:|:|:|:|:|:|
Db 2 HEKLAQLYAVVSEEP 16

RESULT 48
US-09-308-935-6
;; Sequence 6, Application US/09308935
;; Patent No. 6268334
;; GENERAL INFORMATION:
;; APPLICANT: La Thangue, Nicholas B
;; APPLICANT: Bandara, Lisantha R
;; TITLE OF INVENTION: Peptide antagonists of DP transcription factors

;; FILE REFERENCE: 620-67
;; CURRENT APPLICATION NUMBER: US/09/308,935
;; CURRENT FILING DATE: 1999-05-27
;; EARLIER APPLICATION NUMBER: PCT/GB97/03506
;; EARLIER FILING DATE: 1997-12-22
;; EARLIER APPLICATION NUMBER: GB 9626589.7
;; EARLIER FILING DATE: 1996-12-20
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 30
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-6

Query Match 28.9%; Score 28; DB 3; Length 30;
Best Local Similarity 46.2%; Pred. No. 4.2e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSQE 17
:|:|:|:|:|:|
Db 4 LNVLAAMNIIISKE 16

RESULT 49
US-08-436-420-37
;; Sequence 37, Application US/08436420
;; Patent No. 5840524
;; GENERAL INFORMATION:
;; APPLICANT: VAN DAMME, Jo; and
;; APPLICANT: PROOST, Paul
;; TITLE OF INVENTION: GRANULOCYTE CHEMOTACTIC PROTEIN
;; NUMBER OF SEQUENCES: 45
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: HALE and DORR LLP
;; STREET: 1455 PENNSYLVANIA AVENUE, N.W.
;; CITY: WASHINGTON
;; STATE: DISTRICT OF COLUMBIA
;; COUNTRY: USA
;; ZIP: 20004
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/436,420
;; FILING DATE: 24-MAY-1995
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/EP93/03330
;; FILING DATE: 26-NOV-1993
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/982,539
;; FILING DATE: 27-NOV-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BAKER, Hollie L.
;; REGISTRATION NUMBER: 31,321
;; REFERENCE/DOCKET NUMBER: 102378.215
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 942-8400
;; TELEFAX: (202) 942-8484
;; INFORMATION FOR SEQ ID NO: 37:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 36 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-436-420-37

```

Query Match      28.9%; Score 28; DB 2; Length 36;
Best Local Similarity 40.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 4 HLNSKIAFKIVSQEP 18
Db 17 HLKSIQSIKVLSPG 31

RESULT 50
US-09-308-935-1
; Sequence 1, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; EARLIER FILING DATE: 1999-05-27
; EARLIER FILING DATE: 1997-12-22
; EARLIER FILING DATE: 1997-12-22
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-1

Query Match      28.9%; Score 28; DB 3; Length 37;
Best Local Similarity 46.2%; Pred. No. 5.4e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQE 17
Db 11 LNVLMANNIISKE 23

RESULT 51
US-08-905-223-285
; Sequence 285, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655

```

```

; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 285:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -20...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 8.3
; OTHER INFORMATION: seq LLIICHYLPISLC/IP
US-08-905-223-285

Query Match      28.9%; Score 28; DB 3; Length 38;
Best Local Similarity 66.7%; Pred. No. 5.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNHLSN 7
Db 24 PSHINS 29

RESULT 52
US-09-270-767-61791
; Sequence 61791, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61791
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61791

Query Match      28.9%; Score 28; DB 4; Length 41;
Best Local Similarity 41.7%; Pred. No. 6.1e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HLNSKIAFKIVS 15
Db 1 HLNSEEKAKLIT 12

RESULT 53
US-09-205-258-557
; Sequence 557, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER FILING DATE: 1998-06-04
; EARLIER FILING DATE: 1997-06-06
; EARLIER FILING DATE: 1997-06-06
; EARLIER FILING DATE: 1997-06-06

```



```

Db      15 HLRSAILKLSIS 26

RESULT 56
; Sequence 40, Application US/09106568E
; Patent No. 6455248
; GENERAL INFORMATION:
; APPLICANT: Bhattacharjee, J.
; APPLICANT: Suvarna, Kalavati
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DETECTING FUNGAL PATHOGENS IN
; TITLE OF INVENTION: A BIOLOGICAL SAMPLE
; FILE REFERENCE: 96,247-A
; CURRENT APPLICATION NUMBER: US/09/106,568E
; CURRENT FILING DATE: 1998-06-29
; PRIOR APPLICATION NUMBER: 08/650,809
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 40
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Polypeptide segment of ACVS_EMENI shown in Figure 4.
US-09-106-568E-40

Query Match      28.9%; Score 28; DB 4; Length 43;
Best Local Similarity 30.8%; Pred. No. 6.4e+02;
Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy      3 NHLMSKIAFKIVS 15
      ||| : : : :
Db      17 NHLAERLQSEVIS 29

RESULT 57
US-09-270-767-37531
; Sequence 37531, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37531
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-37531

Query Match      28.9%; Score 28; DB 4; Length 44;
Best Local Similarity 62.5%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      9 IAFKIVSQ 16
      | : : :
Db      36 ICFKVVSE 43

RESULT 58
US-09-270-767-52748
; Sequence 52748, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
US-09-270-767-52748

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52748
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-52748

Query Match      28.9%; Score 28; DB 4; Length 44;
Best Local Similarity 62.5%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      9 IAFKIVSQ 16
      | : : :
Db      36 ICFKVVSE 43

RESULT 59
US-09-442-989-30
; Sequence 30, Application US/09442989
; Patent No. 6569993
; GENERAL INFORMATION:
; APPLICANT: Sledeski, Adam W.
; APPLICANT: Mancel, James J.
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF RESIN-BOUND CYCLIC
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: A3113B-US
; CURRENT APPLICATION NUMBER: US/09/442,989
; CURRENT FILING DATE: 1999-11-18
; EARLIER APPLICATION NUMBER: 60/081,897
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)
; OTHER INFORMATION: FMOC-Lys(BOC)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (2)
; OTHER INFORMATION: His(Trt)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (4)
; OTHER INFORMATION: Asn(Trt)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (5)
; OTHER INFORMATION: Ser(OtBu)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (6)..(10)
; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at
; OTHER INFORMATION: position 6 and Asp at position 10 are linked by an
; OTHER INFORMATION: amide bond."
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (7)
; OTHER INFORMATION: Glu(OtBu)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (8)
; OTHER INFORMATION: Arg(PMC)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (6)..(10)
; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at
; OTHER INFORMATION: position 6 and Asp at position 10 are linked by an

```

OTHER INFORMATION: amide bond."
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (11)
OTHER INFORMATION: Trp(BOC)
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (13)
OTHER INFORMATION: Arg(PMC)
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (14)
OTHER INFORMATION: Lys(BOC)
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (17)
OTHER INFORMATION: Gln(Trt)
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (18)
OTHER INFORMATION: Asp(OtBu)
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (6)-(10)
OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at position 6 and Asp at position 10 are linked by an amide bond."
US-09-442-989-30

Query Match 27.8%; Score 27; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 HNSK 8
Db 2 HNSK 6

RESULT 60
US-08-934-915-109
Sequence 109, Application US/08934915
Patent No. 5932412
GENERAL INFORMATION:
APPLICANT: DILLNER, JOAKIM
APPLICANT: DILLNER, LENA
TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN PAPILLOMAVIRUS 1, 5, 6, 8,
TITLE OF INVENTION: 11, 16, 31, 33 AND 56,
TITLE OF INVENTION: USEFUL IN IMMUNOCASSAY FOR DIAGNOSTIC PURPOSES
NUMBER OF SEQUENCES: 193
CORRESPONDENCE ADDRESS:
ADDRESSEE: MASON & ASSOCIATES, P.A.
STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
CITY: CLEARWATER
STATE: FLORIDA
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 3.0
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,915
FILING DATE: 22-SEP-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/949,836
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: LOUISE A. FOUTCH
REGISTRATION NUMBER: 37,133

REFERENCE/DOCKET NUMBER: 1946.6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 813-538-3800
TELEFAX: 813-538-3820
TELEX:
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-934-915-109
Query Match 27.8%; Score 27; DB 2; Length 20;
Best Local Similarity 71.4%; Pred. No. 4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 6 NSKIAFK 12
Db 4 NSEIAYK 10

RESULT 61
US-08-924-629C-12
Sequence 12, Application US/08924629C
Patent No. 6403082
GENERAL INFORMATION:
APPLICANT: Stiles, Michael E.
APPLICANT: Vederas, John C.
APPLICANT: van Belkum, Marius J.
APPLICANT: Worobo, Rodney W.
APPLICANT: Worobo, Rodney J.
APPLICANT: Greer, G. Gordon
APPLICANT: McMullen, Lynn M.
APPLICANT: Leisner, Jorgen J.
APPLICANT: Poon, Alslon
APPLICANT: Franz, Charles M.A.P.
TITLE OF INVENTION: No. 6403082e1Bacteriocins, Transport and Vector System and Methods
FILE REFERENCE: 660.0005US
CURRENT APPLICATION NUMBER: US/08/924,629C
CURRENT FILING DATE: 1997-09-05
PRIOR APPLICATION NUMBER: US 60/026,257
PRIOR FILING DATE: 1996-09-05
NUMBER OF SEQ ID NOS: 80
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 12
LENGTH: 21
TYPE: PET
ORGANISM: Lactococcin A;
US-08-924-629C-12

Query Match 27.8%; Score 27; DB 4; Length 21;
Best Local Similarity 38.5%; Pred. No. 4.2e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Qy 5 LNSKIAFKIVSOE 17
Db 1 MKNQLNFNVSD 13

RESULT 62
US-08-924-629C-80
Sequence 80, Application US/08924629C
Patent No. 6403082
GENERAL INFORMATION:
APPLICANT: Stiles, Michael E.
APPLICANT: Vederas, John C.
APPLICANT: van Belkum, Marius J.
APPLICANT: Worobo, Rodney W.
APPLICANT: Worobo, Rodney J.
APPLICANT: Greer, G. Gordon
APPLICANT: McMullen, Lynn M.
APPLICANT: Leisner, Jorgen J.

```
; APPLICANT: Poon, Alston
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660,000US
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Lactococcin M
US-08-924-629C-80

Query Match      27.8%; Score 27; DB 4; Length 21;
Best Local Similarity 30.8%; Pred. No. 4.2e+02;
Matches 4; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY      5 LNSKIAPKIVSQE 17
Db      1 MKNQLFEILSEE 13

RESULT 63
US-09-228-990-56
; Sequence 56, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 9..13
; OTHER INFORMATION: /product= "OTHER"
; /note= "The side chains of Lys at position 9 and Asp at position 10 are linked by an amide bond."
```

```
; OTHER INFORMATION: 13 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 22
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2."
US-09-228-990-56

Query Match      27.8%; Score 27; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 HLNSK 8
Db      5 HLNSK 9

RESULT 64
US-08-447-411-68
; Sequence 68, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Naja naja
; US-08-447-411-68

Query Match      27.8%; Score 27; DB 1; Length 23;
Best Local Similarity 38.9%; Pred. No. 4.7e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

QY      2 PNH--LNSKIAPKIVSQE 17
Db      5 PTHQDLNLDISIKLPERE 22
```

```

RESULT 65
US-08-447-411-70
; Sequence 70, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; City: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NO. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-447-411-70

Query Match 27.8%; Score 27; DB 1; Length 23;
Best Local Similarity 35.3%; Pred. No. 4.7e+02;
Matches 6; Conservative 5; Mismatches 4; Indels

Qy 2 PNH--LNSKIAFKIVSQ 16
Db 5 PDHKDLNMDVSFHLPSR 21

RESULT 66
US-08-662-227-26
; Sequence 26, Application US/08662227
; Patent No. 592320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA

```


; LENGTH: 23 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-662-227-28

Query Match 27.8%; Score 27; DB 2; Length 23;
 Best Local Similarity 35.3%; Pred. No. 4.7e+02;
 Matches 6; Conservative 5; Mismatches 4; Indels 1; Gaps 2;

Qy 2 PNH--LNSKIAFKIVSQ 16
 Db 5 PDHKDLNMDVSPHLPSPR 21

RESULT 68

US-09-017-947-26
 ; Sequence 26, Application US/09017947
 ; Patent No. 6303754
 ; GENERAL INFORMATION:
 ; APPLICANT: VOGEL, CARL-WILHELM
 ; APPLICANT: BREDEHORST, REINHORST
 ; APPLICANT: KOCK, MICHAEL
 ; APPLICANT: FRITZINGER, DAVID
 ; TITLE OF INVENTION: RECOMBINANT PROCVF
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/017,947
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/662,227
 ; FILING DATE: 14-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 1126-0107-0X
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 23 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-09-017-947-26

Query Match 27.8%; Score 27; DB 3; Length 23;
 Best Local Similarity 38.9%; Pred. No. 4.7e+02;
 Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 2;

Qy 2 PNH--LNSKIAFKIVSQE 17
 Db 5 PTHQDLNLDISIKLPERE 22

RESULT 69

US-09-017-947-28
 ; Sequence 28, Application US/09017947
 ; Patent No. 6303754
 ; GENERAL INFORMATION:
 ; APPLICANT: VOGEL, CARL-WILHELM
 ; APPLICANT: BREDEHORST, REINHORST
 ; APPLICANT: KOCK, MICHAEL
 ; APPLICANT: FRITZINGER, DAVID
 ; TITLE OF INVENTION: RECOMBINANT PROCVF
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/017,947
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/662,227
 ; FILING DATE: 14-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 1126-0107-0X
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 28:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 23 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-09-017-947-28

Query Match 27.8%; Score 27; DB 3; Length 23;
 Best Local Similarity 35.3%; Pred. No. 4.7e+02;
 Matches 6; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

Qy 2 PNH--LNSKIAFKIVSQ 16
 Db 5 PDHKDLNMDVSPHLPSPR 21

RESULT 70

US-09-228-990-57
 ; Sequence 57, Application US/09228990
 ; Patent No. 6472505
 ; GENERAL INFORMATION:
 ; APPLICANT: Condon, Stephen M.
 ; APPLICANT: Morize, Isabelle
 ; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
 ; NUMBER OF SEQUENCES: 88
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, Mailstop 3C43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/228,990
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/017,947
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-OX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-925-442-26
Query Match 27.8%; Score 27; DB 4; Length 23;
Best Local Similarity 38.9%; Pred. No. 4.7e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 2; Gaps 1;
QY 2 ENH--LNSKIAKIVSOE 17
DB 5 PTHQDLNLDISIKLPRE 22
RESULT 72
US-09-925-442-28
Sequence 28, Application US/09925442
Patent No. 6607897
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
BREDEHORST, REINHORST
KOCK, MICHAEL
FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/925,442
FILING DATE: 10-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/017,947
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-OX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/228,990
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,472
FILING DATE: 14-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Martin Esq., Michael B.
REGISTRATION NUMBER: 37,521
REFERENCE/DOCKET NUMBER: A2678B-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-2793
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: Peptide
LOCATION: 10..14
OTHER INFORMATION: /product= "OTHER"
OTHER INFORMATION: /note= "The side chains of Lys at position 10 and Asp at position 14 are linked by an amide bond."
FEATURE:
NAME/KEY: Peptide
LOCATION: 23
OTHER INFORMATION: /product= "OTHER"
OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2."
US-09-228-990-57
Query Match 27.8%; Score 27; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 HLNSK 8
DB 6 HLNSK 10
RESULT 71
US-09-925-442-26
Sequence 26, Application US/09925442
Patent No. 6607897
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
BREDEHORST, REINHORST
KOCK, MICHAEL
FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/925,442

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-925-442-28

Query Match 27.8%; Score 27; DB 4; Length 23;
Best Local Similarity 35.3%; Pred. No. 4.7e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

Qy 2 PNH--LNSKIAPKIVSQ 16
Db 5 PDHKLNDVSHLPSR 21

RESULT 73
US-09-106-568E-111
; Sequence 111, Application US/09106568E
; Patent No. 6455248
; GENERAL INFORMATION:
; APPLICANT: Bhattacherjee, J.
; APPLICANT: Suvarna, Kalavati
; APPLICANT: Bhattacherjee, Vasker
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DETECTING FUNGAL PATHOGENS IN
; TITLE OF INVENTION: A BIOLOGICAL SAMPLE
; FILE REFERENCE: 96,247-A
; CURRENT APPLICATION NUMBER: US/09/106,568E
; PRIOR FILING DATE: 1998-06-29
; PRIOR APPLICATION NUMBER: 08/650,809
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 111
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Polypeptide segment of GRSB_1BACB shown in Figure 4.
US-09-106-568E-111

Query Match 27.8%; Score 27; DB 4; Length 24;
Best Local Similarity 62.5%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HLKSKIAF 11
Db 9 HLKDKFAF 16

RESULT 74
US-09-228-990-58
; Sequence 58, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 11..15
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The side chains of Lys at position 11 and Asp at
; OTHER INFORMATION: position 15 are linked by an amide bond."
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 24
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2."
US-09-228-990-58

Query Match 27.8%; Score 27; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLMSK 8
Db 7 HLMSK 11

RESULT 75
US-09-270-767-57210
; Sequence 57210, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57210
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57210

Query Match 27.8%; Score 27; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NHLNSKIAPKIVSQEP 18
Db 6 NHIKDNELRLPLRQP 21

RESULT 76

```
US-09-270-767-58528
; Sequence 58528, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58528
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-58528

Query Match      27.8%; Score 27; DB 4; Length 24;
Best Local Similarity 55.6%; Pred. No. 4.9e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      7 SKIAFKIVS 15
      |::|::|::|
Db      2 SEVQFRIVS 10

RESULT 77
US-09-228-990-59
; Sequence 59, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 2
; OTHER INFORMATION: /product= "Nle"
```

```
;
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 12..16
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The side chains of Lys at position 12 and Asp at
; OTHER INFORMATION: position 16 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 25
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2.
US-09-228-990-59

Query Match      27.8%; Score 27; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 HLNSK 8
      |::|::|::|
Db      8 HLNSK 12

RESULT 78
US-09-480-993-3
; Sequence 3, Application US/09480993
; Patent No. 6383790
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: High Affinity Kinase Inhibitors for Target Validation
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 51538-5001-US
; CURRENT APPLICATION NUMBER: US/09/480,993
; CURRENT FILING DATE: 2000-01-11
; EARLIER APPLICATION NUMBER: US 60/115,340
; EARLIER FILING DATE: 1999-01-11
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Rous sarcoma virus
US-09-480-993-3

Query Match      27.8%; Score 27; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      11 FKIVSQEP 18
      :::|::|::|
Db      9 YAVVSEEP 16

RESULT 79
US-09-367-065A-3
; Sequence 3, Application US/09367065A
; Patent No. 6390821
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; TITLE OF INVENTION: Nucleotide Triphosphate Substrates
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/367,065A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 26
```

```
; TYPE: PRT
; ORGANISM: Rous sarcoma virus
; FEATURE:
; OTHER INFORMATION: ATP binding domain of v-Src, residues 318-343
US-09-367-065A-3

Query Match      27.8%; Score 27; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      11 FKIVSQEP 18
Db      9 YAVVSEEP 16

RESULT 80
US-09-367-065A-10
; Sequence 10, Application US/09367065A
; Patent No. 6390821
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: Nucleotide Triphosphate Substrates
; CURRENT APPLICATION NUMBER: US/09/367,065A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
US-09-367-065A-10

Query Match      27.8%; Score 27; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      11 FKIVSQEP 18
Db      9 YAVVSEEP 16

RESULT 81
US-09-367-065A-11
; Sequence 11, Application US/09367065A
; Patent No. 6390821
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: Nucleotide Triphosphate Substrates
; CURRENT APPLICATION NUMBER: US/09/367,065A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 26
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
US-09-367-065A-11

Query Match      27.8%; Score 27; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      11 FKIVSQEP 18
Db      9 YAVVSEEP 16

RESULT 82
US-09-367-065A-14
; Sequence 14, Application US/09367065A
; Patent No. 6390821
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: Nucleotide Triphosphate Substrates
; CURRENT APPLICATION NUMBER: US/09/367,065A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
US-09-367-065A-14

Query Match      27.8%; Score 27; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      11 FKIVSQEP 18
Db      9 YAVVSEEP 16

RESULT 83
US-09-367-065A-19
; Sequence 19, Application US/09367065A
; Patent No. 6390821
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: Nucleotide Triphosphate Substrates
; CURRENT APPLICATION NUMBER: US/09/367,065A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
```

```
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: ATP binding domain of Fyn, residues 319-344
US-09-367-065A-19

Query Match          27.8%; Score 27; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      11 FKIVSQEP 18
Db       9 YAVVSEEP 16

RESULT 84
US-09-367-065A-20
; Sequence 20, Application US/09367065A
; Patent No. 6390821
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE OF INVENTION: Nucleotide Triphosphate Substrates
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/367,065A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant Fyn
; OTHER INFORMATION: ATP binding domain, T339G
US-09-367-065A-20

Query Match          27.8%; Score 27; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      11 FKIVSQEP 18
Db       9 YAVVSEEP 16

RESULT 85
US-09-228-990-60
; Sequence 60, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; TITLE OF INVENTION: Morize, Isabelle
; FILE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 3
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 13..17
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The side chains of Lys at position 13 and Asp at
; OTHER INFORMATION: position 17 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 26
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2
US-09-228-990-60

Query Match          27.8%; Score 27; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 HLNSK 8
Db       9 HLNSK 13

RESULT 86
US-09-568-466-3
; Sequence 3, Application US/09568466
; Patent No. 6521417
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE OF INVENTION: Nucleotide Triphosphate Substrates
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/568,466
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 26
; TYPE: PRT
```

```
; ORGANISM: Rous sarcoma virus
; FEATURE:
; OTHER INFORMATION: ATP binding domain of v-Src, residues 318-343
US-09-568-466-3

Query Match      27.8%; Score 27; DB 4; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      11 FKIVSQEP 18
      : : : : :
Db       9 YAVVSEEP 16

RESULT 87
US-09-568-466-10
; Sequence 10, Application US/09568466
; Patent No. 6521417
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/568,466
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
; OTHER INFORMATION: ATP-binding domain, I338G
US-09-568-466-10

Query Match      27.8%; Score 27; DB 4; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      11 FKIVSQEP 18
      : : : : :
Db       9 YAVVSEEP 16

RESULT 88
US-09-568-466-11
; Sequence 11, Application US/09568466
; Patent No. 6521417
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/568,466
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
US-09-568-466-11

Query Match      27.8%; Score 27; DB 4; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      11 FKIVSQEP 18
      : : : : :
Db       9 YAVVSEEP 16

RESULT 89
US-09-568-466-14
; Sequence 14, Application US/09568466
; Patent No. 6521417
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/568,466
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant v-Src
; OTHER INFORMATION: ATP-binding domain, I338S
US-09-568-466-14

Query Match      27.8%; Score 27; DB 4; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      11 FKIVSQEP 18
      : : : : :
Db       9 YAVVSEEP 16

RESULT 90
US-09-568-466-19
; Sequence 19, Application US/09568466
; Patent No. 6521417
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: 51538-5002-US
; CURRENT APPLICATION NUMBER: US/09/568,466
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
```

; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: ATP binding domain of Fyn, residues 319-344
US-09-568-466-19

Query Match 27.8%; Score 27; DB 4; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
: : : : :
DB 9 YAVVSEEP 16

RESULT 91
US-09-568-466-20
; Sequence 20, Application US/09568466
; Patent No. 6521417
; GENERAL INFORMATION:
; APPLICANT: Shokat, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases which Can Utilize Modified
; FILE REFERENCE: Nucleotide Triphosphate Substrates
; CURRENT APPLICATION NUMBER: US/09/568,466
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: US/09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: US 08/797,522
; PRIOR FILING DATE: 1997-02-07
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: PCT/US98/02522
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant Fyn
; OTHER INFORMATION: ATP binding domain, T339G
US-09-568-466-20

Query Match 27.8%; Score 27; DB 4; Length 26;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 FKIVSQEP 18
: : : : :
DB 9 YAVVSEEP 16

RESULT 92
US-08-924-629C-11
; Sequence 11, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon

; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Alston
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082eiBacteriocins, Transport and Vector System and Metho
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Lactococcin A N-terminal;
; FEATURE:
; NAME/KEY: cleavage_site
; LOCATION: (21)..(22)
; OTHER INFORMATION:
US-08-924-629C-11

Query Match 27.8%; Score 27; DB 4; Length 27;
Best Local Similarity 38.5%; Pred. No. 5.6e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSOE 17
: : : : :
DB 1 MNQLNFNIVSDE 13

RESULT 93
US-09-228-990-55
; Sequence 55, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide


```
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 8
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 18..22
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The side chains of Lys at position 18 and Asp at
; OTHER INFORMATION: position 22 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 27
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2."
US-09-228-990-55

Query Match 27.8%; Score 27; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLNSK 8
Db 14 HLNSK 18

RESULT 94
US-09-228-990-61
; Sequence 61, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 4
```

```
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 14..18
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The side chains of Lys at position 14 and Asp at
; OTHER INFORMATION: position 18 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 27
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2."
US-09-228-990-61

Query Match 27.8%; Score 27; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLNSK 8
Db 10 HLNSK 14

RESULT 95
US-09-228-990-54
; Sequence 54, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 8
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 18..22
```

```
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "the side chains of Lys at position 18 and Asp at
; position 22 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 28
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "this C-terminal amino acid is an amide, i.e., CONH2."
US-09-228-990-54
Query Match 27.8%; Score 27; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 5.9e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 HLNK 8
Db 14 HLNK 18

RESULT 96
US-09-228-990-62
; Sequence 62, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 5
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 15..19
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "the side chains of Lys at position 15 and Asp at
; position 19 are linked by an amide bond. "
; FEATURE:
```

```
; NAME/KEY: Peptide
; LOCATION: 28
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2"
US-09-228-990-62
Query Match 27.8%; Score 27; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 HLNK 8
Db 11 HLNK 15

RESULT 97
US-09-228-990-65
; Sequence 65, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 2
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 12..16
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The side chains of Lys at position 12 and Asp at
; position 16 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 28
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2"
```

US-09-228-990-65

Query Match 27.8%; Score 27; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 HLNSK 8
Db 8 HLNSK 12

RESULT 98

US-09-228-990-79
; Sequence 79, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: Peptide
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 2
; OTHER INFORMATION: /product= "Nle"

FEATURE:

NAME/KEY: Peptide
LOCATION: 12..16
OTHER INFORMATION:
OTHER INFORMATION: /product= "OTHER"
OTHER INFORMATION: /note= "The side chains of Lys at position 12 and Asp at position 16 are linked by an amide bond."

FEATURE:

NAME/KEY: Peptide
LOCATION: 20..24
OTHER INFORMATION:
OTHER INFORMATION: /product= "OTHER"
OTHER INFORMATION: /note= "The side chains of Lys at position 20 and Asp at position 24 are linked by an amide bond."

FEATURE:

NAME/KEY: Peptide
LOCATION: 28

; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2."
US-09-228-990-79

Query Match 27.8%; Score 27; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 HLNSK 8
Db 8 HLNSK 12

RESULT 99

US-09-442-989-22
; Sequence 22, Application US/09442989
; Patent No. 6569993
; GENERAL INFORMATION:
; APPLICANT: Sledeski, Adam W.
; APPLICANT: Mencil, James J.
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF RESIN-BOUND CYCLIC PEPTIDES
; FILE REFERENCE: A3113B-US
; CURRENT APPLICATION NUMBER: US/09/442,989
; CURRENT FILING DATE: 1999-11-18
; EARLIER APPLICATION NUMBER: 60/081,897
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 28
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (12)..(16)
; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at position 12 and Asp at position 16 are linked by an amide bond."
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (28)
; OTHER INFORMATION: "This C-terminal amino acid is an amide, i.e., CONH2."
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (2)
; OTHER INFORMATION: Nle
US-09-442-989-22

Query Match

Best Local Similarity 27.8%; Score 27; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 HLNSK 8
Db 8 HLNSK 12

RESULT 100

US-09-442-989-25
; Sequence 25, Application US/09442989
; Patent No. 6569993
; GENERAL INFORMATION:
; APPLICANT: Sledeski, Adam W.
; APPLICANT: Mencil, James J.
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF RESIN-BOUND CYCLIC PEPTIDES
; FILE REFERENCE: A3113B-US
; CURRENT APPLICATION NUMBER: US/09/442,989
; CURRENT FILING DATE: 1999-11-18
; EARLIER APPLICATION NUMBER: 60/081,897
; EARLIER FILING DATE: 1998-04-15

```
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; FEATURES:
; NAME/KEY: Peptide
; LOCATION: 8
; OTHER INFORMATION: /product= "Nle"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 18..22
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The side chains of Lys at position 18 and Asp at
; OTHER INFORMATION: position 22 are linked by an amide bond. "
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 29
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "This C-terminal amino acid is an amide, i.e., CONH2."
; US-09-228-990-53
;
; Query Match 27.8%; Score 27; DB 4; Length 29;
; Best Local Similarity 100.0%; Pred. No. 6.1e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 4 HLNSK 8
; Db 14 HLNSK 18
;
; RESULT 102
; US-09-228-990-63
; Sequence 63, Application US/09228990
; Patent No. 6472505
; GENERAL INFORMATION:
; APPLICANT: Condon, Stephen M.
; APPLICANT: Morize, Isabelle
; TITLE OF INVENTION: PEPTIDE PARATHYROID HORMONE ANALOGS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,990
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,472
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin Esq., Michael B.
; REGISTRATION NUMBER: 37,521
; REFERENCE/DOCKET NUMBER: A2678B-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-2793
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
```